

Let data speak to data

Web tools now allow data sharing and informal debate to take place alongside published papers. But to take full advantage, scientists must embrace a culture of sharing and rethink their vision of databases.

Upload and share your raw data, and have a high impact factor for your blog — or perish? That day has not yet come, but web technologies, from personal publishing tools such as blogs to electronic laboratory notebooks, are pushing the character of the web from that of a large library towards providing a user-driven collaborative workspace (see page 547).

This will in turn expose many fields of research to changes that are already sweeping disciplines such as bioinformatics and high-energy physics. A decade ago, for example, astronomy was still largely about groups keeping observational data proprietary and publishing individual results. Now it is organized around large data sets, with data being shared, coded and made accessible to the whole community. Organized sharing of data within and among smaller and more diverse research communities is more challenging, owing to the plethora of data types and formats.

A key technological shift that could change this is a move away from centralized databases to what are known as 'web services'. These are published interfaces that serve to simplify access to data and software (for an example of such services in action, see www.ebi.ac.uk/xembl/index.html). Until recently the preserve of expert programmers, such interfaces now mean that anyone with even a basic knowledge of programming can automate data processing and analysis.

Various sorts of data are increasingly being stored in formats that computers can understand and manipulate, allowing databases to talk to one another. This enables their users quickly to adapt technologies to extract and interpret data from different sources, and to create entirely new data products and services.

In biodiversity research, for example, rather than creating centralized monolithic databases, scientists could tap into existing databases wherever the data are held, weaving together all the relevant data on a species, from its taxonomy and genetic sequence to its geographical distribution. Such decentralization also helps to solve the problem that databases are often the fruits of individual or

lab research projects that are vulnerable to the vagaries of funding, and to people and labs moving on to pastures new.

Although discipline-specific databases have an indisputable role, science also needs to capitalize on large common repositories for data, whose preservation is guaranteed, and where the data can easily be used by anyone. If that sounds utopian, consider OurMedia, a service created by the Internet Archive and the Creative Commons, which allows anyone to store and share permanently and free of charge any digital work — even their videos and holiday photos. And last month Google launched Google Base, which also allows anyone to upload anything to its massive platform.

Such services will also require new thinking on open data. Web services are dependent on computers being able to freely access data in real time. Although GenBank and many large databases allow unhindered access to their data, many research organizations still cling to obsolete manual data permission policies, which prevent their data being used by web services.

Scientists may be justified in retaining privileged access to data that they have invested heavily in collecting, pending publication — but there are also huge amounts of data that do not need to be kept behind walls. And few organizations seem to be aware that by making their data available under a Creative Commons licence (see <http://creativecommons.org/license>), they can stipulate both rights and credits for the reuse of data, while allowing its uninterrupted access by machines.

As web services empower researchers, the biggest obstacle to fulfilling such visions will be cultural. Scientific competitiveness will always be with us. But developing meaningful credit for those who share their data is essential, to encourage the diversity of means by which researchers can now contribute to the global academy. ■

"By making data available under a Creative Commons licence, scientists can stipulate rights and credits for the reuse of data."

Life at the edge

Successes in structural studies of membrane proteins deserve to be celebrated.

Sealed membrane systems are a defining feature of cellular life. Membranes provide a barrier between the cell and its external environment and, in many organisms, divide the interior of the cell into functionally distinct compartments. The barrier, comprising lipids that are impenetrable to electrically polarized molecules, has proteins inserted within it that allow the selective transport of

ions and molecules. These proteins enable cells to ingest nutrients, excrete metabolic waste, sample the environment for the sake of the immune system, and store energy by means of ion electrochemical gradients. They mediate molecular signalling across the barrier. And they are the very devil to study.

Genome sequencing projects have highlighted the central role of membrane-linked processes in cells. They have revealed that membrane proteins represent about a third of the gene products in most organisms. Unfortunately, our molecular knowledge of these membrane proteins lags far behind that of proteins found in the cell cytoplasm and in external environments. This is primarily due to the difficulty in obtaining high-resolution structural information on



which to build a mechanistic understanding. For example, the purification of membrane proteins for structure determination requires them to be removed from their native membrane environment using detergents. This renders the proteins less stable.

Notwithstanding this and other technical obstacles, isolated successes in the determination of membrane protein structures were reported as early as 1977 for bacteriorhodopsin, the light-powered ion pump in the membranes of archaebacteria, and 1983 for the photosynthetic 'bacterial reaction centre'. However, it is only in the past five years that significant numbers of membrane protein structures have been determined, including structures of ion channels, most components of the mitochondrial and photosynthetic electron transfer chains, and proteins that mediate the transport of small molecules across membranes.

Given this progress, we can now be said to be entering the golden age of membrane protein structure. A flavour of the excitement of membrane proteins can be obtained in many of the articles in the Insight on membrane biology in this issue of *Nature* (see page 577) and by reading the landmark paper on the aquaporin structure embedded in a lipid bilayer (see pages 633 and 569).

But what are the reasons for this recent explosion in membrane protein structures? It is partly driven by technological innovation — the availability of microfocus synchrotron beamlines suitable for data collection from small crystals, advances in the ability to express membrane proteins to high levels and, increasingly, the use of high-throughput screening methods.

In order to obtain crystals one needs to stimulate cells, often of a different organism, to generate, or 'express', large amounts of the required protein. Difficulties in expressing eukaryotic membrane proteins remain the most significant bottleneck in the field. Expressing such proteins in bacterial cells has not been achieved, for reasons that are unclear, and researchers are turning to alternative systems such as insect and yeast cells to obtain their proteins of interest.

But this remains a challenge: milligram quantities are typically required. These challenges in expression, coupled with the difficulty in obtaining diffraction-quality crystals, mean that there is no guarantee of success.

It can take several years, in many cases longer than the standard three-year postdoctoral contract, so embarking on such a quest is a risky business.

Despite the difficulties, there is one area of structural biology where membrane proteins are at the cutting edge: protein structure prediction. Membrane protein structures should be easier to predict than those of water-soluble proteins because the structural possibilities are constrained by the membrane environment.

The prospects for membrane biology are bright, not only thanks to technical breakthroughs but also because of a sense of adventure. The field's success is due in part to the willingness of scientists to dedicate their careers to this challenging endeavour. ■

"The prospects for membrane biology are bright, not only thanks to technical breakthroughs but also because of a sense of adventure."

Stem-cell probe needed

South Korea would benefit from investigating what went wrong in its leading stem-cell lab.

Last week, Woo Suk Hwang of Seoul National University finally admitted using eggs donated by graduate students and paid donors in his embryonic stem-cell research (see page 536). The admission raises pointed questions of the stem-cell research community worldwide and of the South Korean government and media. Each of these groups should be asking themselves why it has taken them so long to take this matter seriously.

In the stem-cell research world beyond South Korea, the adverse publicity generated by Hwang's decision to resign as head of the World Stem Cell Hub should serve as a reminder — as if one was really needed — of the importance of transparent and stringent ethical behaviour by practitioners in this field.

Most of Hwang's international colleagues were slow to accept that anything was amiss in his laboratory. Even as the Korean authorities failed to properly investigate the allegations first made in this journal 18 months ago (*Nature* 429, 3; 2004), researchers seemed almost universally eager to establish fresh collaborations with the laboratory and to laud the quality and integrity of its work.

The withdrawal of collaboration by Gerald Schatten of the University of Pittsburgh on 12 November and Hwang's subsequent confession on 24 November are likely to quench some of that enthusiasm. These events have also opened the doors for South Korea's

media to start looking more closely at the situation, something they were unwilling to do last year.

There are already signs, however, that some people in Korea are drawing the wrong conclusions from the episode. The government has promised to maintain financial support for Hwang, and the organizers of the stem-cell hub say they will refuse his resignation. The Munhwa Broadcasting Corporation, a Korean television network that broadcast a documentary critical of Hwang last week, has been hit by accusations that it is being unpatriotic, and several of the corporation's advertisers have dropped their accounts. On Internet chatboards, producers of the programme have been threatened with violence, and demonstrators have gathered outside the company's Seoul headquarters.

Roh Moo-hyun, the country's president, added his own thoughts to his presidential website on 27 November, calling for calm and branding the actions against the broadcasting corporation "absurd".

Many in Korea still argue that Hwang's ethical lapse was not a serious violation, but merely reflects a difference in culture between Korea and the West. But how can we know what ethical violations — however they are defined — took place, given that there have been so many obstacles to finding the truth? The Korean national interest would best be served not by more flag-waving, but by the rigorous, official inquiry that has yet to be instigated into exactly what went on at Hwang's lab. ■

"How can we know what ethical violations took place, given that there have been so many obstacles to finding the truth?"

RESEARCH HIGHLIGHTS

Expanding horizons

Astron. Astrophys. doi:10.1051/0004-6361:20054185 (2005)

The first results from the Supernova Legacy Survey support the idea that the expansion of the Universe is driven by a 'cosmological constant'. This massive project uses the brightness of supernovae to measure the expansion.

In the 1990s, observations of supernovae revealed that the Universe's expansion was accelerating, pushed by an unknown effect — dubbed 'dark energy' — that counteracts gravity.

Some theoretical models suggest that the density of dark energy can change over time. But the results of this survey, from 71 supernovae, show that this density has been constant to within 10%. By the time the survey ends in 2008, the sample should have swelled to 700 supernovae.



ANGLO-AUSTRALIAN OBSERVATORY/DIMI

IMMUNOLOGY

Taken alive

Immunity **23**, 503-514 (2005)

Immune-system sentinels known as dendritic cells have a reputation for mincing any foreign proteins they meet. They show the remains to T cells, which help to coordinate the immune response. But now US researchers have found that dendritic cells can also capture their prisoners intact and haul them directly before antibody-producing B cells.

Raphael Clynes and his team at Columbia University in New York studied dendritic cells in mice. They found that, with the help of a receptor known as FcγRIIB, cells could swallow a protein and then regurgitate it with its three-dimensional structure intact. Maintaining the protein's structure is key to antibodies recognizing it. What's more, the dendritic cells actively sought out B cells, meaning that B cells do not rely on randomly meeting with microbes, as was once thought.

GENE THERAPY

A great escape

Nature Mater. doi:10.1038/nmat1524 (2005)

Researchers from the University of Tokyo report progress in tackling the biggest challenge in gene therapy — getting DNA into cells safely and effectively. They have developed a light-induced gene-delivery system that minimizes the toxic effects of the treatment by building the DNA and the light-sensitive compound into a single structure.

Large molecules such as DNA tend to be

taken into cells by endocytosis, whereby the outer membrane of the cell pinches off, trapping the complex in a compartment known as an endosome. The photosensitive molecule helps the DNA to escape the endosome by disturbing the membrane when triggered by light. Wrapping the molecule around the DNA minimizes unwanted damage to other organelles and membranes.

The researchers in Japan, led by Kazunori Kataoka, tested their complex in cultured cells and *in vivo* in rats.

QUANTUM OPTICS

As tangled as can be

Phys. Rev. Lett. (in the press);

preprint at www.archive.org/quant-ph/0507128

A pair of photons have been entangled as never before by Paul Kwiat of the University of Illinois at Urbana-Champaign and his team.

When the quantum states of two particles are entangled, you cannot make measurements on one of the pair without determining the state of the other. This is key to quantum information processing, for example in schemes for secure cryptography. The 'hyper-entangled' photons demonstrated by Kwiat's group may offer advantages in some quantum information-coding schemes.

The researchers use two crystals with nonlinear optical properties to generate photon pairs that are entangled in every possible

way: in the quantum states describing their polarization, orbital angular momentum and emission times. In all, this produces entanglement in the 36 different quantum states the photons can adopt.

DEVELOPMENT

Talking about regeneration

Science **310**, 1327-1330 (2005)

Even when sliced up finely, planarian flatworms manage to regenerate themselves from every sliver. Yet they lose this remarkable ability if just one of their genes, *smedwi-2*, is blocked. Researchers have found that blocking the gene prevents the worm's stem cells from maturing into adult cells capable of replacing ageing or missing cells.

The *smedwi-2* gene has counterparts in many species, and the team led by Peter Reddien, at the Whitehead Institute in Cambridge, Massachusetts, concludes that *smedwi-2* may be part of a universal regulatory mechanism for regeneration based on stem cells. In the planarian flatworm pictured below, black regions indicate *smedwi-2* expression.



P. W. REDDIEN & A. SÁNCHEZ-ALVARADO

PHYSIOLOGY

Insects kick up a fuss*J. Exp. Biol.* **208**, 4451–4466 (2005)

A new insecticide, which only affects plant-sucking bugs, was thought to target their central nervous system, as do many insecticides. But it seems that pymetrozine goes after the insects' chordotonal organs, according to Harald Wolf and his team from the University of Ulm in Germany and the Syngenta Crop Protection AG in Switzerland. These organs control the insect's joints, which explains why, in locusts (pictured right), dosed bugs kick up their heels and take on an unusual posture. How this relates to the functional effect of the insecticide — a cessation in feeding behaviour — remains unclear, as does the compound's molecular target.

EVOLUTION

Carping on*Proc. R. Soc. Lond. B* doi:10.1098/rspb.2005.3343 (2005)

Changing body shape to avoid predators can be costly in terms of the ability to compete for food, research on crucian carp (*Carassius carassius*) suggests.

Carp raised in the presence of chemical cues from pike developed a body shape with a greater depth than those growing up without the fear of predation. Similar changes were observed when differing food types were offered, show Jens Andersson and his colleagues at Umeå University in Sweden. Those that were fed on bottom-dwelling chironomids also adopted a deep body shape, whereas those fed on tiny swimming zooplankton had a shallower body.

When offered zooplankton later in the

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

experiment, those carp that had become deeper as a result of either being fed on chironomids or being raised with pike cues were less adept at feeding than their shallow-bodied counterparts.

STEM CELLS

Grow your own liver*Nature Biotech.* doi:10.1038/nbt1163 and 10.1038/nbt1167 (2005)

Making embryonic stem cells turn into liver or pancreas cells should get easier with new methods for producing endoderm — the embryonic germ layer from which liver and pancreas are derived.

Emmanuel Baetge's group at CyThera in San Diego, California, showed that a culture rich in activin A encouraged human embryonic stem cells to differentiate into endoderm, achieving an 80% success rate.

In Japan, a team led by Shin-Ichi Nishikawa of the RIKEN Center for Developmental Biology in Kobe also used activin to induce mouse embryonic stem cells to differentiate into endoderm 25% of the time. Furthermore, they created a monoclonal antibody to monitor the differentiation.

CELL BIOLOGY

Wrinkles in theory of ageing*Cell* **123**, 655–667 (2005)

The protein Sir2 has been shown to prolong life in some species, such as worms and flies, but its action in yeast hints at an undiscovered complexity in the mechanism.

If yeast ageing is measured in terms of how many daughters a cell can produce, Sir2 boosts longevity by 40%. But Valter Longo at the University of Southern California in Los Angeles and his colleagues show that, if the chronological age of the cells is measured, Sir2 limits a yeast's lifespan. The team found that overexpressing the *SIR2* gene in long-lived yeast mutants hastened their demise. They also showed that yeast lacking *SIR2* had reduced rates of DNA mutation and more active stress-resistance genes.

MOLECULAR BIOLOGY

Evasion tactics*J. Exp. Med.* **202**, 1319–1325 (2005)

A compound that protects parasitic eggs from their host's immune system could one day form an anti-inflammatory drug for humans, researchers suggest.

Padraic Fallon of Trinity College in Dublin, Ireland, and his colleagues investigated how the parasitic flatworm *Schistosoma mansoni* evades immune detection. The team knew that certain viruses make 'chemokine-binding' proteins, which bind and neutralize protective immune compounds. The researchers succeeded in isolating this type of protein from *S. mansoni* eggs. They also showed in mice that the protein could suppress skin inflammation and prevent chemokine-induced lung inflammation.

JOURNAL CLUB

Guy Salvesen**Burnham Institute for Medical Research, La Jolla, California**

A cell biologist explains why his colleagues shouldn't look down on cleavages.

If I had a professional mantra it would be, "Don't inhibit the protease that is cleaving your favourite protein; find out what the function of the cleavage is."

Some cell biologists go to great

lengths to avoid proteolysis — or breaking up — of the transcription factors, signalling kinases or cell-surface glycoproteins on which they work. But they could get to love the enzymes that do the snipping, when they realize that the cleavage often defines important aspects of cell signalling.

At my lab we spend a lot of time discovering how cellular processes are governed by proteolysis, and trying to find out how misregulation of proteolytic signalling leads to, for example, degeneration and cancer.

There is now an emerging theme in the field, supported by recent key papers, which suggests that proteolytic activity in the wrong place can change the nature of cells. I have wondered whether a misplaced protease, causing abnormal activation or inactivation of a signalling protein, could promote cell growth and so cause cancer.

Several proteases have been implicated in cancer, but always in a rather roundabout way — through observations of how they influence the progression of tumours

implanted in mice, for example.

So I was intrigued to read about a cell-surface protease that can directly cause skin cancer in mice when its expression is abnormally forced in the epithelium (K. List *et al. Genes Dev.* **19**, 1934–1950; 2005). Because overexpression of the protease in question, matriptase, is correlated with a variety of epithelium-derived tumours in humans, there is finally a chance to understand how mislocation of a protease causes cancer. There is even the prospect of a therapy.

NEWS

Clone star admits lies over eggs

After a year and a half of denials, Woo Suk Hwang admitted last Thursday that his stem-cell research used eggs from paid donors and junior members of his team. Although planning to continue his research, Hwang said he would quit his other positions.

Despite the confession, which shocked South Korea and the global stem-cell research community, many have remained supportive. The Korean government has promised to continue his funding at Seoul National University, and nearly 800 women have signed up to donate eggs to his research through a non-profit foundation. But critics are pushing for a deeper investigation into what happened.

The eggs in question were used in the first successful attempt to derive stem cells from a cloned human embryo (W. S. Hwang *et al.* *Science* **303**, 1669–1674; 2004). The study was hailed worldwide. But three months later, *Nature* reported claims from one of Hwang's graduate students — later retracted — that she and another junior researcher in the lab had donated eggs (see *Nature* **429**, 3; 2004). The controversy was reignited last month, when the fertility doctor who supplied Hwang with eggs, Sung Il Roh, admitted to paying at least 20 women for their donations.

Egg donation is invasive and painful, and can

have serious side effects. Paying donors is now illegal in Korea, although it was not at the time of Hwang's study. And receiving donations from subordinates raises a variety of ethical problems, including the spectre of coercion.

Until 24 November, Hwang had denied everything. Then, in a press conference aired live on all three of the country's main television networks, he admitted that eggs from his researchers and paid donors had been used. A downcast Hwang told the Korean media that he had lied about the researchers' donations to protect their privacy. "I am so ashamed. I will not attempt to justify what I did," he said.

Critics are still concerned, however (see page 532). Young Mo Koo, a bioethicist at Korea's University of Ulsan, says Hwang has not addressed enough questions about his involvement with the egg donors: "There needs to be an investigation by an independent party."

For example, Hwang claims to have known nothing of the payments until a few days before his confession — when Roh told him. Yet in April 2004 he told *Nature* that he had himself "arranged" many of the donors at the hospital concerned. Roh was awarded 40% of the patent resulting from the paper, on which he was not an author. He says he does not know why Hwang offered him so much but that it was not compensation for providing

the eggs. "I don't need any rewards," he says. Hwang has not disclosed his expenditures or budget for the project, saying only that all funds came from private sources.

The extent to which a junior member of his laboratory might have felt pressure to donate is also under debate. The student spoken to by *Nature* last April showed no signs of having been coerced by Hwang. During a 28-minute interview, she proudly described how her patriotism and concern for those with spinal injuries had inspired her to donate. *Nature* was unable to contact the other researcher, who is since thought to have moved to the United States. But according to Roh, she felt obliged to donate after making mistakes early in the experiment that wasted eggs and set the team back by months. "I think it's a beautiful story," Roh told *Nature*, referring to both women's donations.

Insoo Hyun, a bioethicist at Case Western Reserve University in Cleveland, Ohio, has worked with Hwang's group on bioethics protocols. Hyun says the students' accounts prompt hard questions about what coercion is. "To some degree, in Korean society, if you make a mistake you must make good on it somehow," he says. "It's a grey area."

Many Koreans have already sided with Hwang. On 26 November, demonstrators gathered outside the Munhwa Broadcasting Company in Seoul to decry the firm's lack of patriotism after it aired evidence that Hwang

"Hwang has a credibility issue now. The dust hasn't settled."

Antarctic ice puts climate predictions to the test

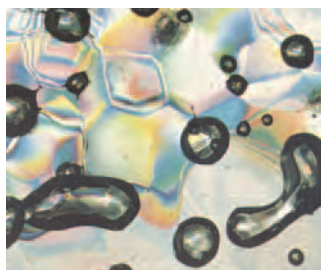
A record of greenhouse gases spanning the past 650,000 years made headlines around the globe last week. The painstaking work proves that levels of carbon dioxide and methane in the atmosphere today massively outstrip those of the pre-industrial era. But it also reveals how little we understand about the way in which these gases influence global climate.

Researchers from the European Project for Ice Coring in Antarctica (EPICA) sampled air bubbles preserved inside a 3,000-metre ice core drilled at Dome C in eastern Antarctica. The data show

previous fluctuations in levels of greenhouse gases in unprecedented detail (see *Science* **310**, 1313–1317; 1317–1321; 2005).

Before the results were in, Eric Wolff, a physical scientist at the British Antarctic Survey in Cambridge, UK, set what came to be known as the 'EPICA challenge'. From previously published data on the prehistoric temperature record, could anyone predict the EPICA gas record? He published the resulting efforts in September (*Eos* **86**, 341–345; 2005).

Now that the actual figures have been published, Wolff says all of the



Bubble whammy: gas in ancient ice shows global models need work.

models correctly predicted that carbon dioxide levels in the period covering the four most recent ice ages would be higher than between

previous ones. The surprise was that entrants who used global models, such as Peter Köhler of the Alfred Wegener Institute in Bremerhaven, Germany, did less well than those who considered only the Southern Ocean, such as Didier Paillard of the Laboratory of Climate and Environmental Sciences in Gif-sur-Yvette, France.

This highlights the degree to which the Southern Ocean influences global greenhouse-gas levels — owing to its size and the fact that its cold, deep waters put large amounts of dissolved carbon dioxide out of circulation. But the less



NOISE RAISES RISK OF HEART ATTACK

High sound levels at home or in the workplace increase heart danger.

www.nature.com/news

J.-M. LEE/AP

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Woo Suk Hwang admits to the world that he used eggs donated by his juniors.

had lied. Eleven out of twelve advertisers have dropped their contracts with the company, according to the *JooAng Daily*.

Although Hwang has promised to resign his directorship of the Seoul-based World Stem Cell Hub, the international cell-sharing initiative says it will hold his position for him.

Some stem-cell researchers and potential collaborators in other countries also seem willing to forgive and forget. One US scientist, who would not speak on the record for fear of

appearing insensitive to ethical issues, says it is hypocritical to punish Hwang for paying donors when infertile couples in the United States often pay tens of thousands of dollars for eggs. "Is it OK to buy them for one purpose and not another?" he asks.

The fact that Hwang lied may make it hard for him to regain his international prestige. "Now it becomes an issue of whether one has a collaborator whose integrity one can trust, and that is a very fundamental issue," says Evan

Snyder, a neuroscientist at the Burnham Institute in La Jolla, California, who has suspended plans to work with Hwang. Hyun agrees: "Hwang has a credibility issue now. The dust hasn't settled."

"It's a useful cautionary tale," adds Larry Goldstein of the University of California, San Diego. He argues that the community should create a set of international ethical guidelines to protect patients and donors.

David Cyranoski and Erika Check

successful performance of Köhler's model, which includes factors such as the effects of vegetation, shows that comprehensive models of the carbon cycle need improvement, says Wolff.

Difficulties in predicting future climate are compounded by the fact that greenhouse-gas levels are set to go off the scale relative to those in the EPICA record. Carbon dioxide did not exceed 290 parts per million in the 650,000 years before the advent of fossil fuels — it now stands at 375 parts per million, and many policy-makers are discussing strategies to stabilize it at an ultimate level of 550. "It seems like a dangerous experiment to me," says Wolff.

The results of that 'experiment'

are already starting to come in. In the past week alone, several groups have warned that the Atlantic Ocean is in trouble (see 'Atlantic feels climate heat'). Researchers hope the findings will focus world leaders, who are currently gathered in Montreal, Canada, to plan emissions policies following the end of the first phase of the Kyoto Protocol in 2012.

"It's clear that we need to stabilize greenhouse gases at a level that allows food production but avoids dangerous interference with the climate," says Thomas Stocker of the University of Bern, Switzerland, who led the latest EPICA analysis. "This cannot be done by quick fixes."

Michael Hopkin

Atlantic feels climate heat

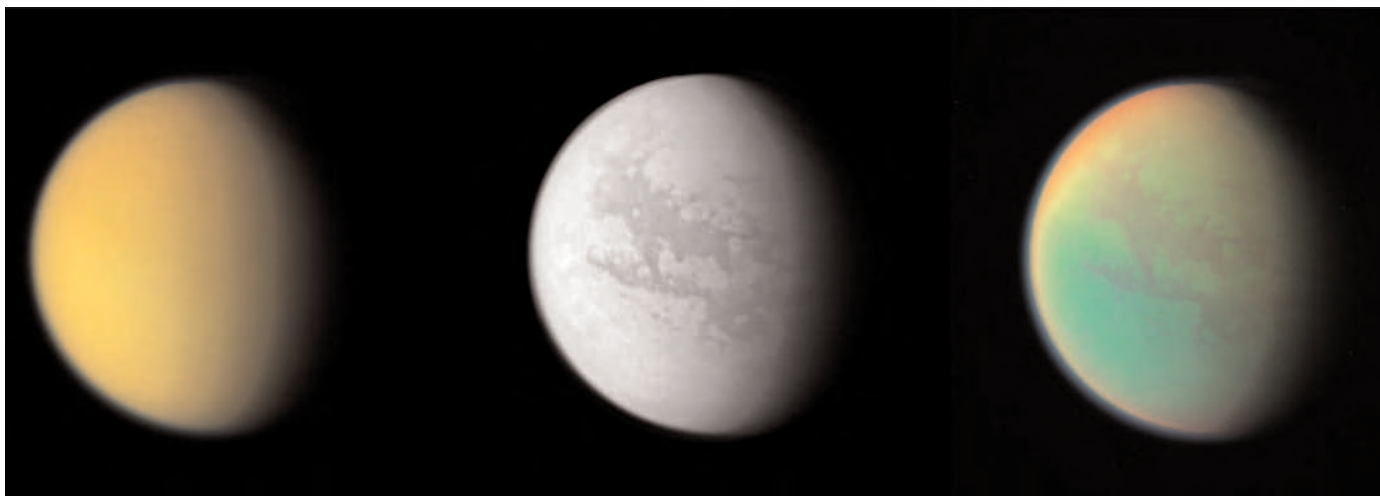
The Atlantic Ocean is already suffering the consequences of global warming, reported climatologists from around the world in the past week. Here are examples of what they have found.

- The North Atlantic is losing its ability to absorb carbon dioxide from the atmosphere. This means that future emissions are more likely to cause global warming, said researchers at a meeting of the European CARBOOCEAN project in Amsterdam.

- The carbon dioxide that does dissolve in the ocean makes it more acidic, threatening to corrode the calcareous exoskeletons of animals such as corals, attendees told the meeting. The Atlantic soaks up some 25% of all carbon dioxide emitted into the atmosphere.

- The system of currents that includes the Gulf Stream — which warms the temperate regions of Europe — is weakening. Research suggests its flow has reduced by a third since 1957 (see page 655). This weakening is evident not in the Gulf Stream itself (the fictional failure of which was dramatized in the film *The Day After Tomorrow*), but in the movement of cold, deep waters.

M.H.



Titan: tapping the flood of data

The first analyses of the Huygens mission to Titan are published this week. **Mark Peplow** charts the satellite's transition from fogbound moon to familiar landscape, and finds out why scientists long to return.

It is easy to forget that just over a year ago, Titan was one of the most mysterious objects in our Solar System. Larger than the planet Mercury, Saturn's largest satellite is the only moon known to have a dense atmosphere. And during years of telescopic observations, Titan has remained hidden beneath its thick hydrocarbon smog, a well-wrapped gift held tantalizingly out of reach.

But since the Cassini-Huygens mission arrived at the ringed planet last year, scientists have been clearing up Titan's mysteries at a tremendous pace. "Before Cassini-Huygens, Titan was the largest unexplored surface in the Solar System," says Ralph Lorenz, a planetary scientist from the University of Arizona, Tucson. "Now scientists are just trying to stand up against the fire hose of information."

Cassini has made nine close fly-bys of the moon, peering through the murk with its radar and infrared sensors, but some of the most exciting details have come from the Huygens lander. This probe hitched a ride halfway across the Solar System with Cassini before descending to Titan's surface on 14 January, and its first results are published online by *Nature* this week (T. Owen *Nature* doi:10.1038/438756a; 2005).

Titan attracts scientists because its methane and nitrogen atmosphere is reminiscent of a primordial, prebiotic Earth. Some hope that understanding Titan's chemistry might reveal how the first molecules of life were formed on our own planet.

The methane was first spotted in 1944 when astronomer Gerard Kuiper studied the

spectrum of Titan's atmosphere. But fleeting visits by the Voyager probes in 1980 and 1981 revealed that the atmosphere was actually mostly nitrogen, squeezed to 1.6 times the Earth's atmospheric pressure at the surface.

Along with a chilly surface temperature of -180°C , this fuelled speculation that liquid hydrocarbon oceans might cover the moon. "If you read the papers beforehand you saw predictions for almost everything imaginable," says John Zarnecki from the Open University, Milton Keynes, UK, who heads Huygens' surface-science team. One of the strongest possibilities was that Titan was the only place in the Solar System besides Earth where rain fell on a solid surface.

Reach for the moon

The only way to find out was to pay a visit. The Cassini mother ship and Huygens probe were built in an unprecedented collaboration between NASA, the European Space Agency and the Italian Space Agency. After a seven-year trek to Saturn, Cassini made its first close pass over Titan on 26 October 2004, and saw fuzzy patterns on the surface that seemed to bear the marks of flowing material.

When Huygens dived through the atmosphere three months later, the pictures became crystal clear, revealing networks of deep channels carved into the landscape. "The striking thing is how much they look like fluvial channels on Earth," says Larry Soderblom of the US Geological Survey in Flagstaff, Arizona. Some channels are 50–100 metres deep, and provide compelling evidence that rain has fallen on the

planet. Other gorges and canyons suggest that lava-like flows and springs have belched liquid from Titan's interior.

This is hard evidence that Titan is geologically active, and has the methane equivalent of a hydrological cycle, with material travelling from surface to skies and back again. "I call it the methanological cycle," says Sushil Atreya, an expert in planetary chemistry at the University of Michigan in Ann Arbor. "Methane really controls the whole atmosphere of Titan."

The probe confirmed that ultraviolet light from the Sun breaks upper-atmosphere methane into fragments, which then combine to form larger organic molecules, such as polyaromatic hydrocarbons. These condense into Titan's characteristic haze of aerosol particles.

Gradually, the aerosols fall from the atmosphere. "Effectively it's grit, snowing down on to the surface," says Soderblom, who is one of the mission's interdisciplinary scientists. According to pictures taken during Huygens' descent, this dark snow collects in river basins and lakes, as though it has been washed off the brighter hills.

Scientists are now certain that the atmosphere is 5% methane: this is not quite enough to produce frequent showers or sustain large lakes, but raises the possibility of occasional monsoons. There is no evidence of oceans, which were thought to be a likely source for the atmosphere's methane — the compound only survives for 10 million to 100 million years in the atmosphere before being broken down.

Carbon isotope measurements also rule out a biological source for the methane — at least

one that is anything like life on Earth. Instead, Atreya suspects that the methane may be generated deep beneath the surface by geological processes that digest carbon dioxide and other organic compounds. A 30-kilometre-wide dome on the surface could be one gate to this underworld — it is thought to be a ‘cryovolcano’, periodically oozing a chilly mixture of water, ammonia and gases from the moon’s innards. And this cold lava could bring up methane from below the surface.

Although the data are impressive, for Zarnecki the highlight of Huygens’ trip is the fact that the probe arrived successfully — it was the most distant controlled landing ever made. “Even though I was part of the team that designed Huygens to be as robust as possible, I was amazed.”

The probe touched down safely on a surface with the consistency of wet sand, its heat throwing up a gentle plume of methane and more complex organic molecules. Images showed Huygens surrounded by smooth pebbles sitting in gullies, apparently hollowed out by flowing liquid. “The surface was a big surprise, it was totally new information,” says Hasso Niemann of NASA’s Goddard Space Flight Center in Greenbelt, Maryland. Niemann, who ran Huygens’ mass spectrometer,

adds that the team still doesn’t know for sure what the pebbles are made of.

Another question unlikely to be solved soon is whether Titan’s chemical soup can make molecules essential for life, such as amino acids; Huygens was not set up to identify chemicals

“Even though I was part of the team that designed Huygens, I was amazed.”

that accurately. The moon’s surface is one of the most promising places in the Solar System for finding these molecules, because it hosts the products of billions of years of photo-

chemistry, which have drifted down from the haze above. Add liquid water and ammonia from cryovolcanoes, and you have an ideal lab for studying how the molecules of life might form over thousands of years. “Nobody can do that experiment on Earth,” says Lorenz. That makes a return visit extremely attractive.

Future forays

As well as being a member of the surface-science team, Lorenz chairs NASA’s Outer Planets Assessment Group, which prioritizes science goals for missions to the outskirts of the Solar System. The group ranks a return to Titan as the second most important mission to work on, after a trip to Jupiter’s ice-covered moon Europa. Titan’s low gravity and thick atmosphere make a robotic airship an attractive prospect, adds Zarnecki, because it could hop around many different sites easily.

Such a mission may take ten years of planning, followed by the seven-year journey to Saturn. But in the meantime Cassini still has at least 36 fly-bys of Titan to complete. Its radar has only mapped a small percentage of the surface, but Soderblom thinks it may be able to cover up to a quarter of the alien landscape. More detail should also come out when Huygens’ and Cassini’s data are compared more closely, and pictures from the descent are tied to radar images from orbit.

“The biggest question is what energy source drives Titan’s surface geology and weather,” says Soderblom. The Sun provides very little light, and radioactive decay and internal movement caused by Titan’s orbit around Saturn don’t seem to add much power.

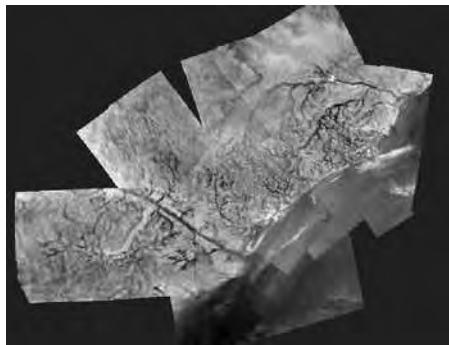
Soderblom suggests that although Titan’s sparsely cratered surface looks young, it could be hundreds of millions of years old. The atmosphere is thick enough to burn up most meteorites before they reach the surface, and drifting dunes of hydrocarbon dust may obscure craters. If it is older than it looks, weak sunlight may be just enough to turn Titan’s methanological cycle in slow motion. There could still be powerful monsoons once every few centuries, points out Lorenz. Any future probes may be well advised to pack an umbrella. ■

ESA/NASA/JPL/UNIV. ARIZONA



Cassini snapped multiple views of Titan (top, left), but these crystal clear images of the surface were Huygens’ reward for plunging below the fog.

ESA/NASA/JPL/UNIV. ARIZONA



Rocky future predicted for labs that rely on postdocs

WASHINGTON DC

Dependence on a temporary and largely foreign workforce is threatening the stability of biomedical research in the United States, a team of biologists warns. If the supply of foreign workers dries up, they say, many labs would not be able to continue functioning as they do.

Susan Gerbi, a biochemist at Brown University in Providence, Rhode Island, and her colleagues tracked the number of researchers in US biomedical labs from 1972 to 2002. They found a dramatic shift away from permanent positions, with temporary postdocs becoming an increasingly vital part of research staff.

In the past, the number of postdocs was roughly equal to the number of principal investigators, but today there are nearly two postdocs for every permanently employed project leader in a lab (see graph). And since 1998, the increase in numbers of postdocs can be wholly accounted for by recruits from abroad. The number of foreign postdocs in biomedical research increased fivefold between 1977 and 2002, the study finds, with temporary US residents now making up more than half of all postdocs.

Gerbi and her colleagues acknowledge that reliance on the brightest workers from overseas over the past few decades has helped the United States to become the world's leader in biomedical research. "This article is not saying that foreign postdocs are bad," says Gerbi.

But they are worried that relying so heavily on foreign workers, especially those with temporary visas, is a dangerous strategy because it masks the relative lack of qualified US

researchers, and means that the whole enterprise could collapse if foreign workers start choosing to go elsewhere. "If left unchanged the situation will deteriorate, and the US scientist will become a dangerously scarce resource," the authors warn (H. H. Garrison, A. L. Stith and S. A. Gerbi *FASEB J.* **19**, 1938–1942; 2005).

Security concerns since the terrorist attacks of 11 September 2001 have made it more difficult for foreign researchers to gain entry to the United States. And countries in Europe and Asia are now actively recruiting foreign talent (see *Nature* **437**, 1215; 2005). "If the wonder-

"If the wonderful foreign postdocs were to dry up, we would be up the creek."

ful foreign postdocs were to dry up because of a political situation," warns Gerbi, "we would be up the creek."

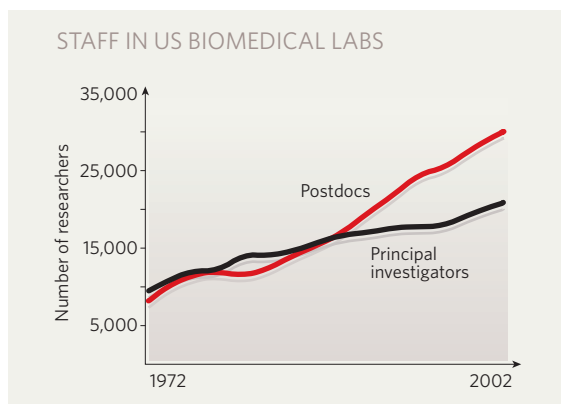
The growing dependence on foreign workers is symptomatic of larger labour problems in US laboratories, according to Alyson Reed, executive director of the National Postdoctoral Association in Washington DC. Postdoctoral fellowships were originally meant as a bridge to more permanent positions, she says. But increasingly the temporary positions are seen by principal investigators as the cheapest way to get highly skilled workers into their labs.

"The principal investigators need to change their ways," says Reed. To create a more stable workforce and encourage home-grown researchers, she says, postdoctoral positions should focus on education, and research labs should employ a higher proportion of permanent staff scientists. For this to happen, Reed admits, principal investigators will need more funding from agencies such as the National Institutes of Health, and more time for mentoring from their universities.

But not everyone agrees on the severity of the situation. "It depends a little bit on your definition of a US scientist," says Richard Freeman, an economist at Harvard University who studies labour trends in higher education. He points out that many foreign-born researchers remain in the United States after their education, and that others continue working for US companies when they return home.

Geoff Brumfiel

SOURCE: FASEB J.



ON THE RECORD

"It's a slap in the face to every Judeo-Christian religion that's out there."

Kansas senator Kay O'Connor (Republican, Olathe) is outraged by a course at the University of Kansas that will teach intelligent design as 'mythology'.

"It's 80 centimetres high with nine leaves, and it looks great."

Sarah Sallon, director of the Natural Medicine Research Center in Jerusalem, describes a date palm that has germinated from a seed 2,000 years old.

Sources: *Lawrence Journal-World*, *National Geographic News*

SCORECARD



British territory

A volcanic eruption has enlarged a remote UK-owned island in the South Atlantic by 0.2 square kilometres.



Paper propulsion

Students at the University of Leeds, UK, have made a paper aeroplane that, in theory, can travel farther than 30 metres.



Dolphin swimming

A study in this week's *BMJ* affirms that spending some time with dolphins can help to dispel depression.

NUMBER CRUNCH

The United Nations aids programme and the World Health Organization have released their latest update on HIV infection.

40.3 million people worldwide are infected with HIV.

4.9 million new infections occurred in 2005.

3.2 million new infections occurred in sub-Saharan Africa this year, although the prevalence of HIV seems to be declining in Kenya, Uganda and Zimbabwe.

270,000 infections occurred in 2005 in eastern Europe and central Asia, the areas experiencing the sharpest increase.

Source: *AIDS Epidemic Update 2005*

Europe's cash crisis puts space plans under threat

Space scientists in Europe are pleading with ministers to significantly boost the European Space Agency's (ESA's) science programme when they meet to decide its budget for the next five years. Without such an increase, researchers fear that high-profile missions may have to be abandoned.

"We are struggling to keep control of the budget," says Marcello Coradini, coordinator of Solar System Missions at ESA headquarters in Paris. Costs of science missions in development are predicted to exceed available money by €300 million (US\$350 million) to €400 million over the coming decade.

If money allocated by ESA's ministerial council, which meets in Berlin on 5–6 December, does not alleviate the pressure on the programme's finances, scientists say it may not be possible to cut back missions

without undermining their scientific goals. And agency officials say they may be reluctant to absorb the costs — roughly equivalent to one year of spending — through delays, because that means no new missions can be started.

The situation has led to speculation that BepiColombo, a mission destined for a 2013 launch to Mercury, might be cancelled. "That is the big danger painted in the sky," says Karl-Heinz Glassmeier, principal investigator on one of the instruments proposed for the spacecraft.

Nerves were set jangling about the project, which also involves the Japan Aerospace Exploration Agency, after it was postponed because the initial design was too heavy. That problem seems to have been solved, but officials say the estimated cost of the mission, at

ESA scientists are concerned that BepiColombo may be cancelled.

€600 million to €650 million, is still more than €100 million above target.

Researchers on BepiColombo are not the only ones worried. "We're all biting our nails,"

Hayabusa ready to head home with asteroid sample

TOKYO

Japan's latest space mission seems to have succeeded in its second attempt to collect pieces of a small asteroid. If so, this will be the first time a sample has been collected for return to Earth from any object in the Solar System apart from the Moon. However, engine trouble casts doubts on whether the craft can return home safely.

On 25 November, the Japan Aerospace Exploration Agency (JAXA) said data sent from Hayabusa show that all stages of the sampling process went well. Agency engineers said it was almost certain that Hayabusa's sampler had touched down on the Itokawa asteroid as planned, and shot two metal pellets into the rock to throw up fragments of the surface. "I think we collected a sample," said project manager Jun'ichi Kawaguchi.

Hayabusa was launched in May 2003, and arrived this September at Itokawa — a potato-shaped, 540-metre-long asteroid located about 300 million kilometres from Earth.

Hayabusa is expected to leave Itokawa by early December for its return journey. Whether a sample was definitely collected will not be known until the craft reaches Earth in the summer of 2007.

To reach this stage, the members of the Hayabusa team have endured a rough time.

Two of the craft's three reaction wheels, which stabilize the probe and help navigation, stopped working; the first failed in July and the second in October. Chemical engines on board were used instead. But their lower accuracy made landing more difficult. In the first landing attempt, on 20 November, Hayabusa seemed to park on the asteroid's hot surface for more than 30 minutes, and failed to collect a sample.

Time and fuel were running short, so the 25 November try was almost the last chance to collect a



Hayabusa snaps its own shadow as it descends.

sample. "It was learning in real time," says Donald Yeomans, US project scientist for Hayabusa and senior research scientist at NASA's Jet Propulsion Laboratory in Pasadena, California. With each attempt, he says, "they learned more and more about how the spacecraft behaves".

Still, concerns remain. Hayabusa lost its balance soon after departing from the asteroid. Engineers are investigating the cause, but one possibility is that the craft's long stay on Itokawa's hot surface during the

first landing attempt damaged one or more of its 12 chemical engines. JAXA said that if the trouble could not be fixed, it would be difficult for Hayabusa to return to Earth.

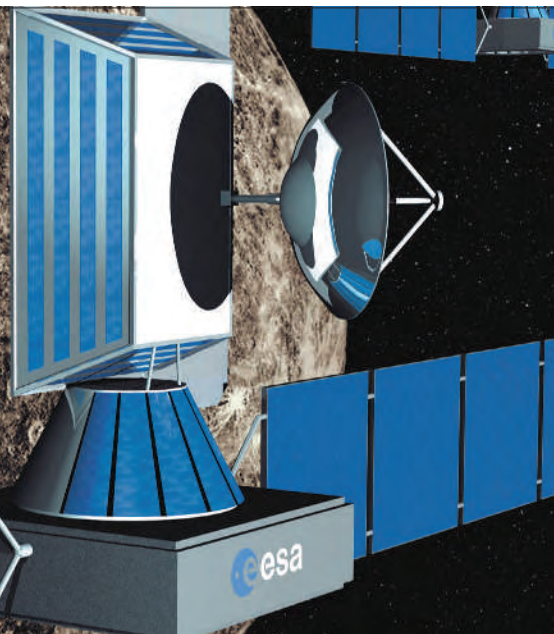
But astronomers have praised Hayabusa's achievements so far as showing the way for future asteroid missions — especially those involving the operation of ion-propulsion engines and delivery of high-resolution images. Analysing the sample, assuming it makes it back to Earth, would also help to answer questions about how the Solar System was created.

The mission is renewing Japan's confidence in space activities. JAXA has recently tried a string of high-risk missions, but has seen many failures over the past few years. "Hayabusa's success has become a tailwind for Japan's space development," Hajime Inoue, JAXA's executive director, said at a press conference. "It proves that the way we have been doing things wasn't wrong." Ichiko Fuyuno

JAXA



THE NATURE PODCAST
Listen to the stories behind the hottest research in our weekly show, available free at www.nature.com/nature/podcast



says Richard Harrison of the Rutherford Appleton Laboratory near Oxford, UK, who works on the Solar Orbiter mission proposed for 2013. "If you look at the list of approved

missions, Solar Orbiter is right at the end. It makes you feel kind of vulnerable."

The reasons for the science programme's money problems are twofold. First, it is still recovering from overruns on the Herschel–Planck mission, which forced the agency to cancel the planet-hunting telescope Eddington in 2003. The Herschel infrared observatory and the Planck satellite, which will measure the cosmic microwave background, are due to be launched together in 2007 and are €178 million over budget.

"It was a big hit," says David Southwood, ESA's director of science. He says the agency is clamping down on missions whose costs spiral, and may cut projects that cannot meet their original budgets. "We are being much more hard-nosed," he adds.

The other factor is that member countries' contributions to the science programme, which makes up 12% of ESA's overall budget, are mandatory, so members must unanimously agree to any increase. The programme's purchasing

power has fallen by 20% since the mid-1990s as payments have failed to match inflation.

Scientists and agency officials are not optimistic about that changing this year, and are urging ministers to reflect on recent successes of the programme. It has 16 operational spacecraft and can boast the notable achievements of Mars Express and Huygens, which touched down on Saturn's moon Titan earlier this year (see page 538), as evidence of their track record. "If we had done very badly, we know we could be financially punished," says Coradini. "But how could we be more successful than we are now?"

He and other officials have asked ministers for a 2.5% year-on-year increase in funding, amounting to about an extra €100 million between 2007 and 2011. That would still leave a shortfall, but Coradini is optimistic: "If we get that amount of money, I think we will not only see the light at the end of the tunnel, we could get out of the tunnel."

Jenny Hogan

"If we had done badly we know we could be punished. But how could we be more successful than now?"

Ireland's science adviser found to have bogus PhD

Barry McSweeney, Ireland's first government science adviser, has stepped down amid allegations that he obtained his doctorate from a degree mill — an unaccredited organization that awards qualifications on the basis of little or no academic study.

McSweeney, a former head of the European Commission's Joint Research Centre who took up his position in June 2004, quit on 15 November after media reports that he obtained his PhD from Pacific Western University in Los Angeles,

California. The university featured in a 2004 US government investigation into degree mills. McSweeney says he was unaware that the university, which granted him a doctorate in 1992, was not accredited.

Announcing his departure, Ireland's minister for enterprise, trade and employment, Micheál Martin, referred to 'Mr' McSweeney, dropping the prefix of 'Dr'. McSweeney will not be out of a job, however: the Irish government has already named him as research coordinator at the Department of Communications, Marine and Natural Resources.

Disputed US stem-cell lines to be tested in Britain

Copies of embryonic stem-cell lines from a large private cell bank in Chicago are to be imported to Britain and made available to the country's researchers.

This month, nearly 150 lines will be brought from Chicago's Reproductive Genetics Institute to the Assisted Reproduction and Gynaecology Centre, a private fertility clinic in London. The quality of cell lines from the Chicago institute, headed by reproductive medic Yury Verlinsky, has been disputed. But few

researchers have worked with them because US federal laws restrict access to lines derived from human embryos.

Mohammed Taranissi, a fertility expert at the London clinic and long-time collaborator of Verlinsky's, will oversee the cell lines. He says that making them available to other researchers will resolve the debate. "I just hope people give this a chance," he says. "There's nothing to lose and a lot to gain."

Germany focuses research on ageing population

Germany's new science minister is to make ageing research a priority. Annette Schavan, a Christian Democrat who took office on 22 November as part of Chancellor Angela Merkel's new government, says millions of euros will be ploughed into research on issues such as degenerative diseases.

Almost 30 million Germans — more than a third of the population — will be over 60 by 2050, and 2 million patients with Alzheimer's disease are expected to need therapy and care. Schavan wants to increase investment in areas such as stem-cell research — mainly involving adult stem cells as the law that restricts research using human embryonic stem cells will remain unchanged, she says.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Moving on: Barry McSweeney has quit his post as science adviser to the Irish government.

Responsibilities for science have also been rearranged. Most basic research, including life sciences and biotechnology, will remain under the aegis of the research ministry, but aeronautics, space and maritime research will be handled by the ministry of economics.

Researcher found guilty of negligent homicide

A cancer researcher has been sentenced to nearly six years in prison for the negligent killing of a patient enrolled in a drug study.

Paul Kornak of the Stratton Veterans Affairs Medical Center in Albany, New York, admitted to forging medical reports in order to make patients eligible for drug studies that he and a colleague undertook between 1999 and 2002. Among the patients was 71-year-old James DiGeorgio, an Air Force veteran who died a few weeks after participating in a stomach-cancer study.

The experimental chemotherapy that DiGeorgio received was never explicitly linked to his death, but Kornak has admitted to forging blood tests so that the patient could participate. He pleaded guilty on 18 January to the charge of criminally negligent homicide.

Burglary warning closes London diamond exhibition

In the case of the Natural History Museum in London, diamonds are not forever.

The 'Diamonds' exhibition there was closed permanently without notice on 23 November after London police told the museum's director that they believed criminals were planning to target it.

The exhibition, which explains the science behind the formation of diamonds,



had been scheduled to run until February. On show were many of the world's most valuable diamonds including the Millennium Star (shown here).

This is not the first time criminals have targeted an exhibition containing the Millennium Star. Police disguised as cleaners foiled an attempt in 2000 to steal the diamond from a display at the Millennium Dome in London.

Radar provides deeper view of ice on Mars

Mars Express, the European Space Agency's orbiting probe, has provided a glimpse deep below the surface of Mars.

According to data from the MARSIS radar instrument, the layer of water ice at the planet's north pole is about 1.8 kilometres deep and is almost pure (*Science* doi:10.1126/science.1122165; 2005). Contrary to the hopes of some researchers, the probe did not find any melted water

between the ice and the sandy material that lies beneath it. The possibility of finding water had excited researchers, as such an environment could potentially harbour martian organisms.

MARSIS has also found what appears to be an underground impact basin, up to 250 kilometres across, in the Chryse Planitia region. Scientists hope that further radar soundings could reveal more hidden craters. The experiments had been delayed for a year because of fears that unfolding the MARSIS antennas could destabilize Mars Express.

DE BEERS



The expanding electronic universe

Spoofs, like scientists, need to keep an eye on the future. So when a veteran researcher at the CIA extols the virtues of the latest tools on the Internet, it is worth tapping in to his thoughts. Calvin Andrus, head of the CIA's unit for collaboration technologies, set out his stall in September's *Studies in Intelligence* with a paper entitled "The wiki and the blog: toward a complex adaptive intelligence community". Intelligence officers, he argues, should have access both to online (although obviously restricted) blogs on which they can record their experiences and insights, and to wikis — websites that can be edited by a community. Such tools could transform the responsiveness of the intelligence community, helping isolated officers to comment on each other's ideas and to collate rapidly breaking data and information.

This week's News Features look at what wikis, blogs and other technologies may mean for the future of scientific communication beyond the confines of scientific journals. These tools offer fresh opportunities both before publication, when people are debating ideas and hypotheses, and after, when they are finding and discussing published results. They also provide scientists with exciting new possibilities for communicating with policy-makers and the public.

Our opening Feature on scientific blogs and wikis finds that scientists are lagging behind other communities, including the commercial sector, in seizing these opportunities. Young scientists are often reluctant to express their thoughts online out of fear that it is somehow inappropriate, or even possibly damaging to their careers.

Another worry is that new technologies bring the threat of

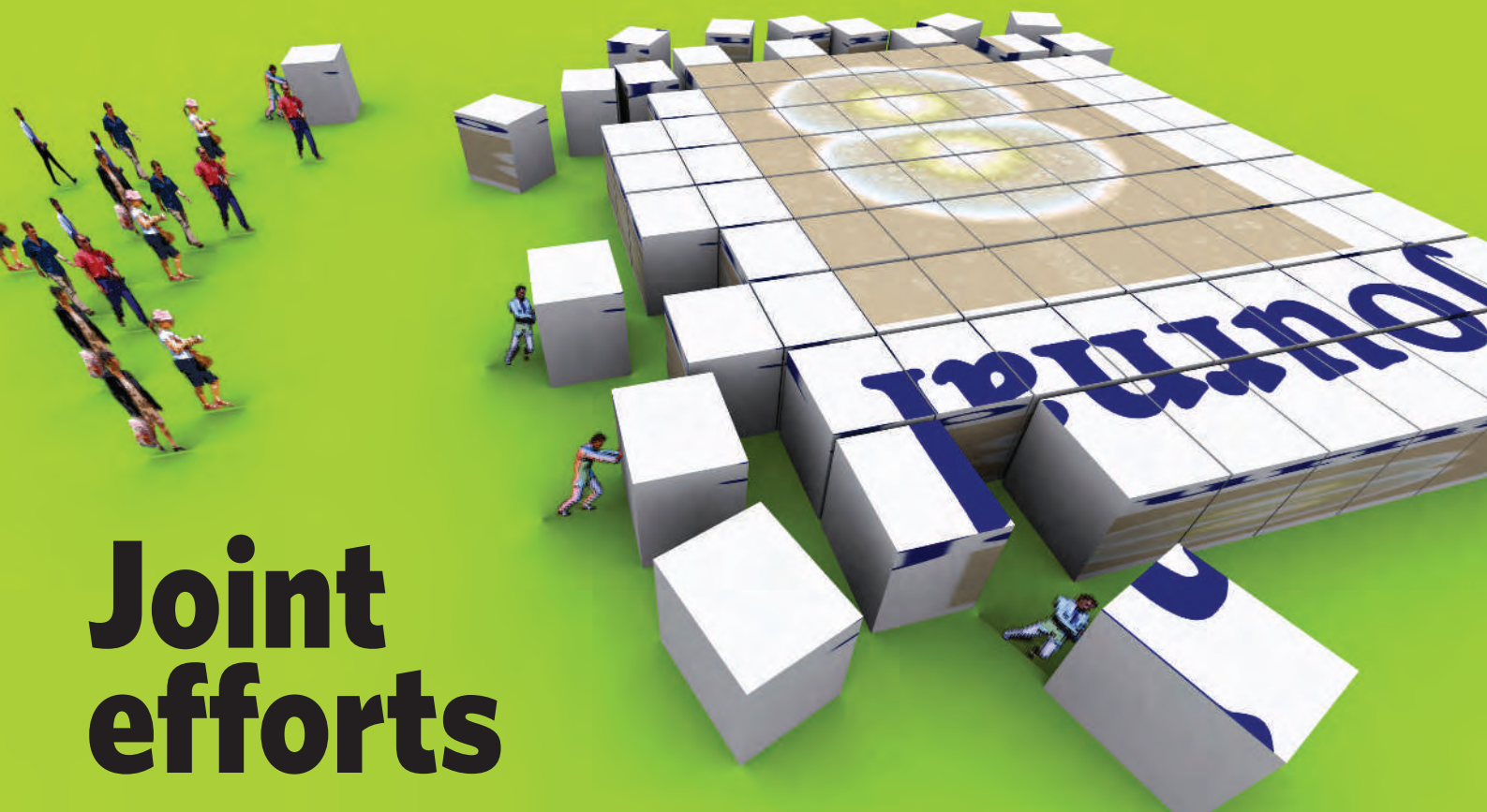
information overload. Some of the tools we highlight here offer hope on this score, whether through better filtering or improved searching. Academic librarians are keeping a close eye on the development of search engines, such as Google Scholar; some are even blogging about it (<http://acrlblog.org>). But the technological idea that is most likely to land them in hot water is the growth of digital libraries.

University libraries that have signed up to digital scanning projects, such as Google Book Search, are finding themselves in the uncomfortable position of leading the revolution. This suggests that they have a vision for their future, something readers who want to discover new works and read books in new ways will welcome. But they may have some explaining to do to publishers, including their own university presses, and to authors who are gearing up for a legal battle with the search giant.

A fair number of these possibilities will no doubt fail to bear fruit, either nobly — by losing out to even better ideas and technologies — or ignobly, held back by the scientific establishment's cultural resistance. Some, however, will surely thrive. As Andrus puts it: "For every ninety-nine mediocre ideas, there will likely only be one brilliant idea. The few brilliant ideas, however, are worth the investment of many mediocre (and chaotic) ones... The few brilliant ideas will survive in the market place of ideas."

Sarah Tomlin is News Feature and Commentary editor.

For online versions of these News Features with active links, please go to www.nature.com/news



Joint efforts

At its best, academia is a marketplace of ideas. But many scientists are reluctant to embrace the latest web tools that would allow them to communicate their ideas in new ways, says **Declan Butler**.

When Tim Berners-Lee invented the World Wide Web in 1989, he saw it as a collaborative workspace for his fellow scientists at CERN, the European particle-physics lab near Geneva, and beyond. His creation went on to surpass his prediction that "the usefulness of the scheme would in turn encourage its increased use". But in the rush to develop the web as a flexible way to find information, the original concept of users interacting in real time was largely forgotten. Fifteen years later, the web seems to be returning to its roots.

For most users, the web in its first decade was like a big online library, where they mainly searched for information. Today it is undergoing a subtle but profound shift, dubbed Web 2.0, to become more of a social web, not unlike Berners-Lee's original vision. Yet scientists are largely being left behind in this second revolution, as they are proving slow to adopt many of the latest technologies that could help them communicate online more rapidly and collaboratively than they do now.

"I find it ironic that science is about the adoption, discovery and exploitation of new knowledge and techniques, yet the biggest revolution on the web is passing us by," says Greg Tyrelle, a bioinformatician at Chang Guan University in Taiwan. He has been experimenting with blog (short for web log) software for five years to interact with a growing audience of his peers and the wider public.

The emerging web is largely being shaped by dynamic interactions between users in real time. But many researchers still see publications in the formal scientific literature as 'the' means of scientific communication. Although the traditional published paper is accepted as the undisputed information of record, younger researchers, in particular, are concerned that scientists are missing out on new ways to communicate with each other and the public.

They recommend the use of collaborative technologies such as blogs and wikis, websites that any visitor can add to and edit. Supporters say these offer a forum for broader and more

timely discussion, to complement the existing system of peer-reviewed journals. This could enhance science communication, both before publication, when generating ideas, and after publication, when discussing results (see 'Open house', opposite).

Blogs are just one example of new social technologies that are allowing more people to publish more easily and in more diverse ways on the web. By allowing reader feedback and syndication feeds, blogs create an instant online community. "Blogs can offer any kind of content — from peer-reviewed articles to sheer speculation to rants, and everything in between," says Amy Gahran, an expert in new media and editor of Contentious.com.

The write stuff

The best-known wiki is the online encyclopaedia, Wikipedia, which has grown to almost a million entries since its launch in 2001. Scientists at Harvard and the Massachusetts Institute of Technology (MIT) recently started their own wiki, OpenWetWare, to apply the same approach to sharing lab protocols and data among biology groups worldwide.

Outside academia, blogs are taking off in a big way. A study published in October by the Guidewire Group, a research firm in new media, says that 90% of marketing communication companies have either launched, or intend to launch, internal blogs. There are now some 20 million blogs, permeating almost every sector of society. But science is a glaring exception, and today there are still only a few dozen scientific bloggers.

"Until blogging is seen as normal, worries about what your supervisors think will continue to be a problem."
— Gavin Schmidt



C. DANKIN

C. FIELD

Scientists who blog see their activities as a useful adjunct to formal journals, not a replacement. "The standard scientific paper is irreplaceable as a fixed, archivable document that defines a checkpoint in a body of work, but it's static, it's very limited," says Paul Myers, a biologist at the University of Minnesota, who blogs at Pharyngula.

"Put a description of your paper on a weblog, though, and something very different happens," says Myers. "People who are very far afield from your usual circle start thinking about the subject. They bring up interesting perspectives." By sharing ideas online, you get feedback and new research ideas, he says.

A senior US epidemiologist who blogs once or twice a day under the pseudonym 'Revere' on his public-health blog Effect Measure, has attracted a diverse readership. "About 1,500 people visit each day," he says. "If someone told me that I could show up at a lecture hall every day and deliver a short opinion, and that 1,500 people would show up to hear me, I'd be pretty satisfied — 1,500 is twice the subscription of many speciality journals."

But for most scientists and academics, blogs and wikis remain unattractive distractions from their real work. Many consider them an online version of coffee-room chatter, background noise that goes against the very ethos of heavily filtered scholarly information.

Opinion pieces

Scientists who frequent the 'blogosphere' see it differently. The dynamic hierarchy of links and recommendations generated by blogs creates powerful collaborative filtering, they argue. Blogs may create noise, but they are a great way of keeping up with what's hot in your field, says Tyrelle, who blogs at Nodalpoint.org. He believes that the more bloggers there are in a particular community, the more efficient this filtering becomes, so — counter-intuitively — reducing information overload.

Tyrelle suggests that this is not so different from BioMed Central's Faculty of 1,000, a popular fee-based service that highlights biology papers according to recommendations from a subset of 1,000 scientists. But in the blogosphere, this service is free and could marshal input from a subset of 10,000 scientists or more.

Yet even the most web-savvy scientists remain unconvinced that blogs have any useful role in science. "I have my doubts that blogging reduces information overload, but blogging will survive as it appeals to all the exhibitionists," quips Rolf Apweiler, a bioinformatician at the European Bioinformatics Institute in Hinxton, UK, and head of the UniProtKB/Swiss-Prot protein-sequence database.

Others disagree. "Science is too hung up on the notion of 'the paper' as the exclusive means of scientific communication," says Leigh Dodds, a web expert at the publisher Ingenta. Publication and research assessments are more geared to measuring a researcher's standing than communicating science, he claims.

Open house

Online pioneers they are not, but traditional publishers are not entirely stuck in the past. Publishing online often means bundling supplementary information with a mirror copy of the print article, but the web is now being used to open up some journals to more interactive discussions — previously only possible at conferences.

The *BMJ* website led the way in allowing readers to post 'rapid responses' to published articles. But in June this year, the *BMJ* changed its criteria for accepting online contributions — adding heavier moderation. Journals thinking of adding companion blogs (see main text) will also want to moderate comments.

Atmospheric Chemistry and Physics (ACP), published by Copernicus, uses online discussion to open up the peer-review process. Papers

are published online quickly and referees post comments online, anonymously if they wish. Authors, and other researchers, can chip in as long as they identify themselves. After the discussion is closed, editors use it to shape the final version of a paper.

Advocates say the online debate improves the final product. "It lets others see what the leading people in the area are thinking and forces editors, referees and authors to work at a higher standard," says Scot Martin, an environmental chemist at Harvard University and an editor at *ACP*.

Arne Richter, managing director at Copernicus, has high hopes for the journal, which has gained a healthy impact factor of 2.7 since its 2001 launch. But Richter admits that of six Copernicus journals with online discussion, not all have been welcomed by users. *Hydrology*

and *Earth System Sciences* added open peer review seven years after its launch. "A tribe of very conservative scientists keeps asking why there has to be a discussion feature," says Richter. "They just don't want it."

The editors of a new online journal to be published by BioMed Central think biologists are ready for open peer review. *Biology Direct* authors have to solicit their own reviews from an editorial board, and the comments appear online for all to see.

"In many areas of biology there's roughly a one-in-three chance one of your reviewers just won't like your point of view," says editor-in-chief David Lipman. If that were to happen to a *Biology Direct* paper, it would still be published. But anyone could read the naysayer's comment.

Tom Simonite

Jennifer Hallinan, a biologist at the University of Queensland, Australia, who runs the blog Cancer Dynamics, agrees with him. The web is providing a hierarchy of sources, she says, including useful blogs and wikis. "Each level of the hierarchy has its own sources of error, its own strengths and weaknesses," she explains, "but these are known and can be taken into account when using them."

Blogs associated with traditional journals may help bridge the gap between the literature and blogs, says Glenn McGee, editor-in-chief of *The American Journal of Bioethics*. The leading journal in its field, it was the first to create a companion blog, *Blog.Bioethics.Net*.

"Put a description of your paper on a blog, and people far from your usual circle start thinking about the subject."

— Paul Myers



The bioethics blog allows the journal to respond faster and in different ways to public controversies, says McGee. The blog has high impact, he adds, often influencing reporting on ethical issues by the mainstream media.

Print journals cannot keep up with developments in certain fields, adds Gavin Schmidt, a researcher at NASA's Goddard Institute for Space Studies in New York, who blogs at Real-

Climate.org with other climate scientists. The blog helps to reduce noise by setting the record straight, says Michael Mann, another RealClimate blogger and director of Pennsylvania State University's Earth System Science Center, citing as an example a recent post on whether hurricanes are linked to global warming (see www.realclimate.org/index.php?p=181).

McGee and Schmidt have permanent jobs, and both agree that many scientists don't blog because they fear it has a poor image and could damage their careers. Most younger biologists blog anonymously, says Roland Krause, a researcher at the Max Planck Institute for Molecular Genetics in Berlin and a bioinformatics blogger. "Many fear that their superiors consider it a waste of time, or even dangerous," he says. Schmidt agrees: "Until blogging is seen as normal, this will continue to be a problem."

Others fear being scooped by rivals. "In many institutes it's just way too dangerous to discuss work in progress with the people across the floor," regrets Krause — let alone on a blog.

Such fears are dated, argues Jason Kelly, an MIT graduate student involved in OpenWetWare. The upcoming generation, he says, believes that excessive competition can harm science; they see the benefits of brainstorming their research ideas on blogs as far outweighing the risks.

Kelly admits some may regard this view as naive. But Schmidt suggests that once scientists come up with some sort of peer-review mechanism for blogs that increase their credibility, without diminishing their spontaneity, blogs will take off.

Declan Butler is a senior reporter at Nature.

THE REAL DEATH OF PRINT

Despite clashes with publishers over copyright, Google's plan to make millions of books available online is turning the tide for efforts to digitize the world's literature.

Andreas von Bubnoff tracks the demise of the printed page.

Vishwas Chavan travels a lot. An informatician based at the National Chemical Laboratory in Pune, India, he collects data on what types of animal live where in India to enter into a biodiversity database. Yet the specimens he hunts for have neither fur nor feathers, but yellowing pages and ageing dustjackets.

Much of the information Chavan seeks is in old, out-of-print tomes that are scattered around the world; about 2,500 of the 7,000 books he has unearthed were written in the first half of the nineteenth century. To find them, Chavan has spent years trailing around libraries. He dreams of the day when books such as these are scanned and made available as digital files on the Internet.

Chavan and other digitization visionaries paint a future in which books no longer gather dust on shelves, but exist as interconnected nodes in a vast web of stored literature, all accessible at the click of a mouse. So instead of hunting for specific books, scholars could search for specific information, customizing searches to suit their needs.

A few years ago, Chavan's dream seemed little more than a castle in the air. True, a number of mostly volunteer-driven or publicly funded projects had been scanning books and making them freely available on the Internet. But most efforts were limited.

In December 2004, the Internet search-engine company Google announced plans to change that. It said it would scan millions of books from five major libraries: the university libraries of Oxford, Harvard, Stanford and Michigan, and the New York Public Library. The announcement energized other organizations in the United States and in Europe, which soon unveiled similar plans to

scan and catalogue millions of books.

The move to digitize books is set to transform the worlds of publishers, librarians, authors, readers and researchers. Obscure specialist titles could find new readerships; librarians and information specialists will have to develop tools to catalogue and navigate this labyrinth of data; and authors and publishers may soon have to start thinking in digital dimensions, just as website designers and writers already do.

Bloody revolution

But revolutions are rarely bloodless and this one could soon get ugly. In the United States authors and publishers are squaring up against Google for a legal fight over copyright. Opinion is divided over whether the scanning projects being implemented by companies such as Google and Amazon (see graphic opposite) will hand control of the world's literature to private enterprise — and, if so, what this could mean. And with several independent scanning projects under way, it is still not clear how much of the information will be freely avail-

able, or where and how it can all be coordinated and accessed.

The idea to digitize books and make them available online has been around since the Internet's inception in the early 1970s. When the US Declaration of Independence was typed in and sent to everyone on a computer network on the night of 4 July 1971, it marked the birth of Project Gutenberg, the first book-digitization venture.

Since then, the project's 20,000 volunteers have scanned or typed in about 50,000 out-of-copyright books, says its founder Michael Hart, who works in the basement of his home in Urbana, Illinois, and, like the project's volunteers, for free.

Projects such as this are driven by the idealistic desire to make knowledge and literature freely accessible to all, but also by the benefits of having book collections easily searchable. "Being able to find it online is pretty much the same as having it online," says David Weinberger of the Berkman Center for Internet and Society at Harvard Law School in Cambridge, Massachusetts.



C. DANKIN

Assets such as searchability have prompted the National Science Foundation (NSF) in Arlington, Virginia, to get involved in an open-access enterprise called the Million Book Project. This is an international scanning effort with many participants, including Carnegie Mellon University in Pittsburgh, Pennsylvania.

Since the project began in 2002, about 600,000 out-of-copyright books have been scanned, although only about half of them are currently available online (see graphic). The scanning takes place in India and China, with books being shipped there temporarily from libraries around the world.

Made to fit

Searchability is also the main driving force behind commercial plans to scan books, including texts whose copyright has yet to expire. For example, if their products have been digitized, online booksellers can allow customers to search within books and browse a few pages before deciding to buy. In the United States, with the publisher's permission, Amazon puts searchable digital data from mostly copyrighted books online. Amazon says that several hundred thousand books are currently available for searching.

Amazon also offers the option of purchasing e-books and e-documents on its website, which can be viewed after downloading them to a portable reading device (see 'Will flexible screens be the end of paperbacks?'). The company expects these services to drive additional sales. Its 'search inside the book' feature increases sales by 8%, the company says. Scientific publishers, such as the US National Academies Press also see increased print sales when they allow their books to be viewed online.

But Google doesn't mention money in its announcement that it plans to make the contents of millions of copyrighted books searchable as part of its Google Book Search project. Its spokesman, Nate Tyler, says Google's motivation is to include literature that is currently only available offline in its mission to make information universally accessible. But the possibility that the company could gain financially from the move has raised hackles among US authors and publishing organizations.

"Google is in a class by itself because of the quantity of money and the level of centralization."

— Daniel Greenstein

In the autumn of this year, the Authors Guild and the Association of American Publishers filed a lawsuit against Google for copyright infringement. They complained that Google hadn't asked them for permission to scan copyrighted books.

Google has obtained the go-ahead from publishers to include some copyrighted works as part of its Book Search project, but not all. It argues that it does not need to seek permission for every book, because what it plans to do is permissible according to the 'fair use' exception of US copyright law. This allows copying for uses such as teaching, scholarship or research.

Google will, for example, not make the full text available, but only show 'snippets' of text around the search results if a book is still copyrighted. The company says that people are more likely to buy or borrow a book if they can search it this way, adding that the snippets are similar to the card catalogues found in

libraries. But Paul Aiken of the Authors Guild in New York City argues that the act of scanning the works is copyright infringement no matter how the texts are used.

The outcome of the lawsuit will depend on the courts' decisions over how the concept of fair use applies in the age of digital books and the Internet. Meanwhile, the rest of the scanning world is watching from the sidelines, and being careful to scan only books that are out of copyright, or to obtain the publisher's permission before scanning anything.

Google's plan has shaken up the digital-book world in other ways too. For one thing, many believe that its size and resources mean Google can pull off this feat — so large-scale repositories of digital books seem a more realistic and immediate prospect than ever before. Google has also galvanized its competitors, both public and private (see graphic) to redouble their efforts, and has placed a question mark over the future of libraries and librarians.

"I think Google is in a class by itself because of the quantity of money and the level of centralization," says Daniel Greenstein, librarian of the California Digital Library in Oakland, California. "Google has paved the way, created the appetite for this kind of activity, and anxiety on the part of libraries and publishers."

Out with the old

But Michael Gorman, president of the American Library Association, says he is not worried that libraries could become obsolete. As well as providing access to books, they serve as a place for people to meet and study, he says. And librarians' expertise in information management will still be needed. "We are not worried about our own jobs," agrees Dennis Dillon,



C. DARKIN/N. SPENCER

Will flexible screens be the end of paperbacks?

"It's nice. It's the size of a paperback book. It's very light — I actually forgot I had it in my bag."

The future of reading, according to Theresa Horner, director of e-book operations for the publisher Harper Collins, lies in neat electronic devices such as the Sony Librie she totes in her handbag. Although some bookworms might balk at the thought of reading a novel on a screen, it might not be long before portable electronic books revolutionize the book world in the same way that Apple's iPod changed the music scene.

Almost any device can be used to read an e-book. Nick Bogaty, executive director of the International Digital Publishing Forum (formerly the Open e-Book Forum) in New York thinks that the "vast majority" of people reading e-books now are doing so on their PDAs or smartphones. The small, handheld devices that people use to organize their lives can act as useful stores for reference books, or hold novels that while away a long commute. Fans of this approach say they stop noticing the screen size

when gripped by a good plot.

There are also devices that attempt to recreate a more familiar reading experience. Garth Conboy, president of eBook Technologies, recounts how his company's devices, which are like small tablet-PCs with simple page-turning buttons, are supplied to about 15,000 students at a military college loaded with their course material. This saves the college

money in printing, and has the advantages of online material.

But the device that aficionados drool over is the Sony Librie. Launched in Japan in April last year, it is similar in size, shape and weight to a book, and is the first device to have a paper-like display.

The display uses technology developed by E Ink in Cambridge, Massachusetts, and works by reflecting light, rather than emitting

light like an LCD or LED screen. The pixels are microcapsules containing particles of black and white ink with opposite charge. An electric field pushes one or other colour to the surface, producing a print-like effect. The device only gobbles energy when the display is refreshed. So the lifetime of batteries in such an e-paper system is measured in page-turns, rather than hours.

And the future? It's likely to be flexible. In September, Philips unveiled the first working prototype of an e-reader with a rollable display, aimed not just at e-books, but also at news reading and Internet browsing. In the 'Concept Radius' a grey-scale screen measuring about 13 centimetres wide, again using E Ink technology, curls up inside a slimline packet measuring 10 × 6 × 2 cm.

It might be worth waiting for. "We plan to start our production by the end of 2006, so I anticipate products coming into the market in early 2007," says Hans Driessen, a spokesman for Philips.

Jenny Hogan

S. KAMBAYASHI/AP

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

associate director of the research services division of the University of Texas libraries at Austin. "The job is changing, which makes it even more fulfilling than it was before."

But Gorman is worried that over-reliance on digital texts could change the way people read — and not for the better. He calls it the "atomization of knowledge". Google searches retrieve snippets and Gorman worries that people who confine their reading to these short paragraphs could miss out on the deeper understanding that can be conveyed by longer, narrative prose. Dillon agrees that people use e-books in the same way that they use web pages: dipping in and out of the content.

Best of both worlds

Having a mix of both e-books and real books could be the answer. A mix would certainly help solve that perennial headache for libraries — the lack of shelf space and cost of keeping physical books. Ensuring that some libraries always keep a physical copy of a particular work means that they will be available through inter-library loans for readers needing a real book, adds Dillon.

Some of them are already dispensing with hard copies. The University of Texas at Austin, for example has about 10,000 copyrighted books and 300,000 out-of-copyright works that are available only as e-books, says Dillon.

Another person to be energized, but also alarmed, by Google's move is Brewster Kahle,

founder of the Internet Archive, a non-profit organization in San Francisco that archives web pages and other digital files. Although Google has never indicated that it plans to claim ownership over its digitized material or charge for search access, Kahle doesn't want to leave digital books entirely in the hands of private enterprise.

That's why, in October, he announced the formation of the Open Content Alliance (OCA). This aims to build a permanent

"Science is moving incredibly fast, and scanning old books is a complete waste of money."

— Matthias Ulmer

archive of multilingual digitized text and multimedia content, which, as far as possible, will be freely accessible.

Like the Million Book Project, the OCA will scan out-of-print books; the first few are already available online. The alliance hopes to rival Google's project in terms of scale; among the groups helping to finance its scanning efforts are Yahoo and Microsoft. Some libraries who were reluctant to join the Million Book Project for logistical reasons have signed up to the OCA. "We did consider the Million Book Project, but we were hesitant because we wanted to avoid shipping overseas," says Tom

Garnett, assistant director for Digital Libraries and Information Systems at the Smithsonian Institution Libraries in Washington DC, a contributor to the OCA.

For taxonomists such as Chavan, the OCA is perhaps the most interesting scanning project so far. Eight museums including the Natural History Museum in London have formed the Biodiversity Heritage Library Project, which will collaborate with the OCA to scan about one million volumes of biodiversity literature, much of it out of copyright.

But Matthias Ulmer, a German publisher who helped launch an e-book initiative by the German Publishers and Booksellers Association, thinks that scanning old books is "a complete waste of money". "Science is moving incredibly fast, even in the field of taxonomy," says Ulmer. Earlier this year, his association announced an initiative whereby some 100 German publishers are considering digitizing about 100,000 newly published books by 2006. Publishers will take their own digital raw data and place them on a network of their own servers. Scientists and others will then be able to access the books for a fee.

With 2005 seeing the birth of so many digitization projects, it might not be long before Chavan can realize his dream of hunting for new specimens from the comfort of his armchair.

Andreas von Bubnoff is an intern in Nature's Washington office.

Start your engines

Google has launched another challenge to commercial search services — this time aimed at scientists. But is the new engine running as smoothly as its fans hope? **Jim Giles** investigates.

As an undergraduate in India in the mid-1980s, Anurag Acharya had to write letters to scientists when he could not find the papers he wanted. It is a memory that makes the softly spoken computer engineer laugh. Now working at Google, Acharya is creating a search tool that aims to be the first choice for everyone from Indian students to Iranian professors. "I want to make it the one place to go to for scholarly information across all languages and disciplines," he says. And that ambition, he freely admits, is "simple to state, but not to achieve".

For a member of the public seeking a one-off scholarly article, Google Scholar is ideal. It is free to access, and as easy to use as the main Google search engine (see 'Inside information', opposite). But for academics with access to dedicated library resources, why make the switch? Most scientists rely on tried and trusted favourites, including subject-specific databases such as the US National Institutes of Health's PubMed or the NASA Astrophysics Data System, to find papers.

Since its launch last November, Acharya's Scholar engine has delighted and infuriated in equal measure. One librarian has even begun a blog following the search engine's progress. Although there are no detailed studies, many librarians report that faculty members and students are beginning to use the search engine; some suspect that Scholar will replace more established, and more costly, search tools. Figures from academic publishers also suggest that use of Scholar is growing rapidly: it already directs more online traffic to *Nature* websites than any other multidisciplinary science search engine.

Thomas Mrsic-Flogel, a neuroscientist at the Max Plank Institute of Neurobiology in Martinsried, Germany, and a regular PubMed user, has started to use Scholar. He says he finds the engine useful when he is not quite sure what he is searching for. Search results include citation links to other articles, so he follows the links until he finds something interesting — a function that PubMed, which does not track citations, cannot provide. "I follow the citation trail and get to papers I hadn't expected," says Mrsic-Flogel. "I have found papers that way that I wouldn't have found otherwise."

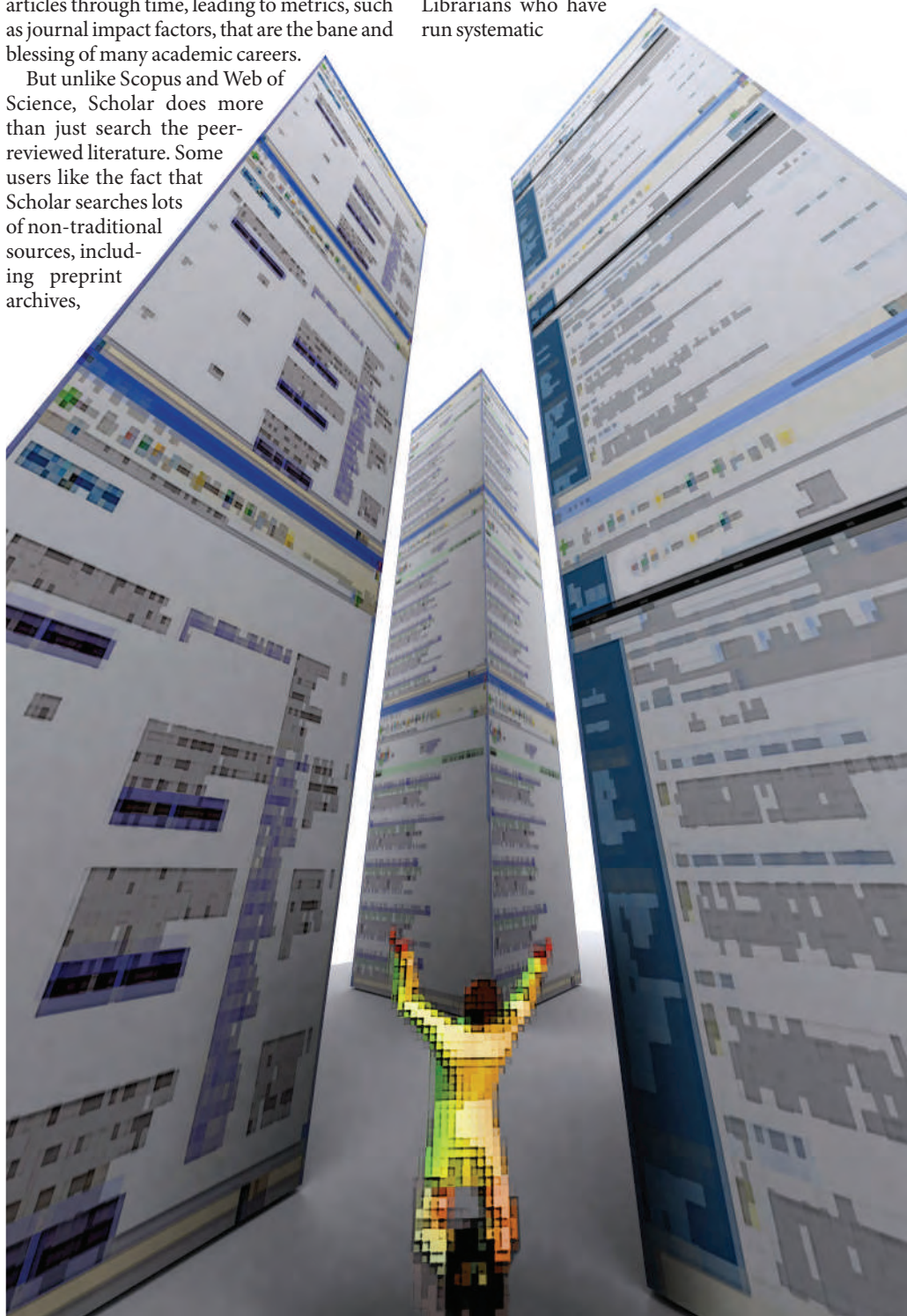
This citation tracking puts Scholar in direct competition with the fee-based search engines marketed by traditional science publishers. Until Elsevier launched its search engine, Scopus, in 2004, Thomson Scientific's Web of

Science had a monopoly on citation tracking. Citation counts allow researchers, institutes and journals to follow the impact of individual articles through time, leading to metrics, such as journal impact factors, that are the bane and blessing of many academic careers.

But unlike Scopus and Web of Science, Scholar does more than just search the peer-reviewed literature. Some users like the fact that Scholar searches lots of non-traditional sources, including preprint archives,

conference proceedings and institutional repositories, often locating free versions of articles on author websites. This 'grey literature' is growing in importance but remains poorly defined. It is widely assumed that Google considers a source scholarly if it is cited by another scholarly resource — but as online publishing evolves, so may this definition. Advocates of greater access to the scientific literature hope that Scholar will encourage more researchers to deposit their articles in free online repositories.

But how well does Scholar actually work? Librarians who have run systematic



Inside information

Science search engines are fine for literature searches, but scientists inevitably need much broader information from the web. Searching using the main Google engine may take some coaxing, but a few tricks can help you to find the most relevant information faster, and to get a variety of views on a topic.

Google has advanced search options that will help you narrow your search, using more precise terms, or broaden it, using synonyms. Here we list some less well-known tips, using the drug Tamiflu as an example.

Site: Websites are often difficult to find your way around, so rather than wasting time endlessly clicking, just type 'site:' into your query followed by the website name. Searches can also be restricted to a domain name. For example, 'site:gov' will limit a search to US government sites, and 'site.nih.gov' to the National Institutes of Health. A search for Tamiflu at the World Health Organization, 'Tamiflu site:who.int', returns about 100 hits. A broader search, such as 'tamiflu site:edu', brings back more than 40,000 hits from US universities.

Filetype: A useful way to refine searches is to search for particular document types using the 'filetype:' query. A search for 'Tamiflu filetype:ppt' will return only PowerPoint presentations, which are usually conference talks. 'Filetype:doc' will often return project proposals or government texts, 'filetype:pdf' is more likely to return scientific information.

Define: This simple query will provide a definition of the words you enter after it, gathered from various online sources. The query

'define:Tamiflu' takes you to definitions in Wikipedia in several languages for example.

Quotation marks Ultimately, the web is about people, and if you are looking for contacts, or possible collaborators, there are some ways to Google scientists. The query, "'avian influenza" "workshop participants"', will bring back a few hundred hits, often with contact details for world experts among the top results. Variations of this will do the same in any scientific field.

Declan Butler

searches across several engines, say that Scholar performs well. A study published this year, which looked at more than 100 papers, concludes that Scholar finds similar numbers of citations to its commercial rivals¹. Yet such results need to be interpreted cautiously, say information scientists. Critics point out that the study did not examine the list of citations to see whether they contained duplicated or erroneous entries.

A closer look at Scholar search results suggests that duplication may well be occurring. One of Scholar's harshest critics, Péter Jascó, an information scientist at the University of Hawaii in Honolulu, has taken the engine on numerous test drives. He has documented the results in unflattering terms on a website run by Thomson Scientific. In one extreme case, Jascó found that the first 100 results from a search for documents on 'computers' and 'intractability' returned 92 slightly different citations of a book entitled *Computers and Intractability* and only 8 other unique results.

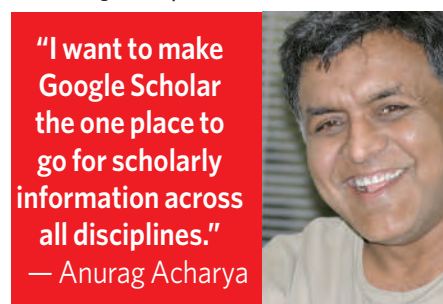
Cite unseen

The source of this problem is the way in which Google adds records to its scholarly index. At Web of Science and Scopus, staff scan in the abstracts and references from print journals and use dedicated electronic feeds supplied by publishers. Scholar, by contrast, uses an automated process. Software robots crawl the web in search of documents that look like scientific papers, and then use algorithms to strip out relevant information such as author and publication date. The process is vastly cheaper and quicker, but it is not yet updated daily and there are no manual checks to delete duplicates or correct misclassified records.

Google has deals with several academic publishers that allow it to search the full text of many papers, whereas Web of Science and the others are largely restricted to searching abstracts. But Scholar's index is restricted to online sources — Web of Science has archives that go back to 1900. And the automated

process means Scholar's citation tracking can return odd results. For example, Web of Science finds almost 14,000 citations for a 1988 *Science* paper on the polymerase chain reaction², identifying it as the most highly cited paper ever to appear in that journal. Scholar finds just under 3,000.

All this suggests that there may be little overlap between the citations in the grey literature found by Scholar, and those extracted from the primary literature — even when the



citation counts match up. For now, librarians are unanimous in their advice: stick to Web of Science or Scopus if you need to do a thorough literature search or an accurate citation count. The engines have impressive coverage and well indexed records with fewer misclassified entries. Librarians also warn that Scholar is still an experimental, or beta, version. Google remains reluctant to reveal details of its search algorithm, or what it indexes, so hopes of using Scholar as a tool for checking on citation counts is a distant prospect, they say.

All three search engines will continue to evolve. Scopus and Web of Science plan to add additional resources to their databases, such as institutional repositories, together with new ways for searching those sources. Scopus, for example, is integrated with a chemical database, such that users can go from a literature search to see structural information on molecules of interest. But it is unlikely that these engines will ever mine the grey literature as broadly as Scholar. Elsevier has a separate, free search engine, called Scirus, that searches science web

resources, but it doesn't track citations.

So where does this leave Acharya's bold goal? Librarians say that Scholar's current high usage rates are likely to reflect searches run by undergraduates, who typically require only a couple of key papers on any one subject, and researchers who want a quick snapshot of an unfamiliar field. Acharya says he intended Scholar to appeal to such users, but also wants to attract academics who need to keep up with the latest papers in their field. As Thomson and Elsevier continue to invest in new services, it will be interesting to see whether Scholar can keep up.

Two's company

With just two full-time staff working with Acharya, it would seem that Scholar is a low priority for Google. But maybe they could draw on the expertise of outside computer programmers by letting them write software that taps into Scholar's database. It is an approach Google has used before to good effect. If Google allows programmers to do the same with Scholar, it is likely that additions would be developed by librarians and academics.

So does Scholar plan to open itself to outsiders? Not right now, says Acharya. He remains cagey, but is not ruling it out. "We may reconsider this decision once the service is closer to how we envisage it."

The Google team may also reconsider if enthusiasm for Scholar continues to grow. Librarians at Virginia Tech in Blacksburg have already created a free software extension, called LibX, for an Internet browser, which allows users to retrieve papers using Scholar with a simple mouse click on highlighted text. LibX will take you directly to your library's resources, if the paper can be found there. And that is the sort of tool both Google and librarians can learn to love.

Jim Giles is a reporter for Nature in London.

1. Bauer, K. & Bakkalbasi, N. *D-Lib Magazine* 10.1045/september2005-bauer (2005).

2. Saiki, R. K. et al. *Science* **239**, 487–491 (1988).

BUSINESS

Swiss star finds it tough at the top

Serono is calling in outside advice. Colin Macilwain investigates the future of Europe's leading biotech firm.

Large-scale success has eluded most European biotechnology companies. But on the banks of Lake Geneva stands solid evidence that it can indeed be achieved.

There, at a cost of SFr340 million (US\$260 million), Serono is building a new headquarters. From next autumn the complex will house 1,200 research and administrative staff, and it provides concrete proof that the Swiss company is in the top tier of global biotechnology.

For Serono's 600-or-so early-stage researchers, the company's environment strikes a happy balance between the innovation of a small start-up and the stability of a major corporation. "We sit in the sweet spot between the two cultures," says Tim Wells, Serono's head of research. "We live with cash flows that are luxurious compared with those of a small, start-up company."

But last month, the future of that balance began to look precarious. On 8 November, Serono confirmed that it had called in investment bank Goldman Sachs to explore 'strategic options' for the company. It was also reported that several major drug companies have been in discussions about the possibility of buying the firm. As is normal in such situations, Serono officials declined to discuss the corporation's plans; and the intentions of the Bertarelli family — which owns two-thirds of the company's stock — remain unknown.

For sale or not, Serono is a prominent feature on Europe's rather barren biotech landscape. With almost 5,000 employees and annual revenues of US\$2.5 billion, it is by far the largest biotechnology operation on the continent. In terms of sales, it is the third largest in the world, after California's Amgen and Genentech.

Deep roots

Unusually for a biotech firm, the Swiss company has deep historical roots. It was founded 99 years ago in Rome by pharmacologist Cesare Serono. Initially it extracted hormones for fertility treatment from urine, but it managed to make the smooth transition into producing much more profitable, genetically engineered versions of these compounds.

Today Serono holds more than half of the global market in fertility treatments with drugs such as GONAL-f (follitropin- α), which



Serono scientists have benefited from an innovative and diverse research environment in Geneva.

induces ovulation in infertile women. But two-thirds of its revenues are generated by Rebif (recombinant β -interferon), which is used to treat multiple sclerosis.

Wells says that the company has succeeded in establishing an innovative environment in Geneva, where it does about two-thirds of its research. "There are many nationalities," he says, "and the sort of multiculturalism that empowers US biotech companies."

Serono has established an extensive network of arrangements with European research universities, including major collaborations in neuroscience with the University of Zurich and in multiple sclerosis with the San Raffaele Scientific Institute in Milan, Italy.

Most of the company's remaining scientists work in Boston where, in common with so many others in the industry, Serono's labs are trying to develop cancer treatments.

But despite the strong cash flow from its established products and its success in creating a first-class research set-up, Serono has experienced a string of recent disappointments with late-stage products in its pipeline. About half of the 13 new compounds it was developing in 2003 have hit trouble, including potential treatments for psoriasis and, in partnership with CancerVax of California, for melanoma. "They've really been relying on one product," says Peter Knight, who watches the biotech

industry for Edinburgh-based market analysts Wood Mackenzie. "Like everyone else, they need to feed the pipeline — and they've been pretty unsuccessful at doing it."

That may have led to the arrival of Goldman Sachs and the discussions with prospective buyers. At the head of the queue is Pfizer, which already distributes Rebif in the crucial US market. "There are a number of large pharmaceutical companies who have so far steered clear of biologics who could see Serono as a way in," says Keith Redpath, vice-president of life sciences at Wood Mackenzie, noting that such a buyer would view the firm "as a biotech research and development operation that can wash its face financially".

There is still no certainty that such a sale will take place. Analysts have questioned whether Serono is worth a probable purchase price of US\$15 billion, given that it has no assured winners in its drug pipeline. Some say that a partnership with another biotech that has prospects in different areas of medicine might make more sense.

And if there is a change of ownership at Serono, outside observers say that what would matter to the European biotech sector as a whole is that the company's research and development operation remains intact. That way, it can continue to provide a model for other European biotechs to aspire to — and, with a bit of luck, unearth some more successful drug candidates.

"Serono needs to feed the pipeline — and it's been pretty unsuccessful at doing it."

Automated grading of research performance clearly fails to measure up

SIR — More contentious than national rankings of research quality, as shown, for example, by David A. King (*Nature* **430**; 311–316; 2004), is the application of such measures to research institutions. Several leading organizations — such as Thomson Scientific (formerly Thomson ISI), the centres for science and technology studies in Leiden (CWTS) and Bern (CEST) and European bibliometric analysts (see A. F. J. van Raan *Scientometrics* **62**, 133–143; 2005) — have emphasized the risk of reaching erroneous conclusions through using inappropriate data.

We have compared an automated and a manual analysis of the performance, between 1994 and 2003 of two European national organizations: the UK Medical Research Council (MRC) and the French biomedical research agency (INSERM) in France. The two agencies are both devoted to biomedical research and are of comparable size.

We first used Thomson Scientific's Web of Science, which correctly identified all 17,829 publications from the MRC and all 46,978 from INSERM. We then compared the Essential Science Indicators (ESI) from the Thomson ranking with a manually extracted list of the 'top 1%' of publications affiliated to France and Britain.

The results turn out to be very different. The manual analysis took affiliations into account carefully, whereas the automated index missed many INSERM-affiliated papers. ESI rankings show 253 'top 1%' publications for the MRC and 117 for INSERM, whereas the manual count has the two organizations on a more equal footing, with 513 top 1% publications for the MRC and 535 for INSERM. As many as 50% of the MRC's and 80% of INSERM's highly cited publications are not identified by the automatic extraction.

Given the use to which these figures are put by funding agencies and governments, these discrepancies, and discrepancies in other types of citation studies, emphasize the problems that can arise from the use of bibliometric analyses.

It is important to ensure that affiliations are captured correctly before performing an analysis, and to use the appropriate citation measure.

For both the MRC and INSERM, only about 20% of papers published in high-impact journals are in the 'highly cited' category, demonstrating that the two indicators should not be confounded.

The research organization of France is extremely complex, which renders assessment difficult. But we believe that

France and other countries must collaborate and reach agreement on benchmarks for assessment of research performance, including a simplified, generally accepted affiliation nomenclature.

N. Haeffner-Cavaillon, C. Graillot-Gak,

C. Bréchet

Cellule de Bibliométrie, Département de l'Évaluation Scientifique, INSERM, 101 rue de Tolbiac, 75654 Paris Cedex 13, France

Animal-rights zealots put wildlife welfare at risk

SIR — Your Editorial calling for government resistance to intimidation from animal-rights lobbyists ("Taking a stand on animal-rights violence" *Nature* **438**, 1; 2005) provides timely advice for researchers too, as our public scientific meetings are increasingly attracting disruptive protest if they involve animal research.

In October I participated in the annual scientific meeting of the Royal Zoological Society of New South Wales. The focus was on the challenges of managing the impacts of over-abundant animals and pest species that threaten Australia's biodiversity and economy. These meetings are open to the public, and discussion is encouraged.

But the 2005 forum was systematically and strategically sabotaged by animal-rights lobbyists from at least six organizations. They monopolized question time and plenary discussion sessions with prepared speeches and interjections, all pushing a short-sighted single-issue agenda.

They were preaching to the converted about the fundamental need for animal welfare, as all who attended the forum share a concern for the protection of Australia's unique wildlife and environment. And by strangling discussion and learning among the very people charged with looking after Australia's wildlife and agriculture, these lobbyists did not help their primary cause: reducing the killing of animals by humans.

Many novel ideas on how to manage pest impact while minimizing the need for direct control were entirely lost in a flood of rhetoric that all animals should live. Yet this right to life apparently does not include those animals that are maimed, killed, displaced or even driven to extinction by the over-abundant pests that continue to plague Australia. Such animals were overlooked by these zealots with their 'let it be' approach to conservation.

This was a great opportunity lost. It sends a grim warning to other scientific societies and researchers hoping to exchange ideas in open forums about animal-related issues.

Peter B. Banks

School of Biological, Earth and Environmental Sciences, University of New South Wales, Kensington, New South Wales 2052, Australia

Evaluation bias hits women who aren't twice as good

SIR — In your Editorial "All things equal" (*Nature* **437**, 296; 2005) and Special Report "Small steps towards campus child care" (*Nature* **437**, 446–447; 2005), much was made of the need for women scientists to have access to good child care if they are to succeed. However, this recent attention to child care in the scientific workplace merely addresses a symptom, rather than a cause, of under-representation of women in science.

Childless women and those with children have strikingly similar patterns of salary disparity and lag in achieving tenure and promotion compared with men.

As your report highlights, nations differ in child-care facilities — but they all share a shortage of women scientists, particularly at higher levels. Furthermore, the proportion of women in different sub-disciplines varies dramatically, but child-care availability is no different for a microbiologist or an engineer.

We suggest that lying behind the paucity of women in science is an unconscious bias in evaluating the sexes. Research shows that both men and women tend to overrate men and underrate women in competence, particularly when women are in a non-traditional field such as science (V. Valian *Why So Slow?* MIT Press, Cambridge, MA, 1998). For example, when the heads of 147 psychology departments were sent fictitious resumé of prospective faculty members and asked to name the rank — assistant, associate or full professor — to which the candidate would be appointed in their department, the recommended rank was higher if the resumé had a male name than if the same qualifications had a woman's name attached (L. S. Fidell in *Woman: Dependent or Independent Variable?* 774–782, eds R. K. Unger and F. L. Denmark, Psychological Dimensions, New York, 1975).

More recently, women had to produce twice as many scientific papers of equivalent quality as men to be considered equally competent in a Swedish Medical Research Fellowship postdoctoral programme (C. Wennerås and A. Wold *Nature* **387**, 341–343; 1997).

Marlene Zuk*, Gunilla Rosenqvist†

*Department of Biology, University of California, Riverside, California 92521, USA

†Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

Contributions to Correspondence may be submitted to corres@nature.com. They should be no longer than 500 words, and ideally shorter. They should be signed by no more than three authors; preferably by one. Published contributions are edited.

BOOKS & ARTS

Different class

The 'big three' universities in the United States are upholding a long tradition of élitism.

The Chosen: The Hidden History of Admissions and Exclusion at Harvard, Yale, and Princeton

by Jerome Karabel

Houghton Mifflin: 2005. 672 pp. \$28

John Aubrey Douglass

In the decades after the American War of Independence in the late eighteenth century, many of the fledgling nation's political leaders envisaged a collection of colleges and universities as central agents of maturation for those born in the glow of the Enlightenment. In a society dominated politically by farmers, shopkeepers and artisans, the concept of the university as a source of intellectual development, civic leadership and socioeconomic mobility contrasted strongly with the stifling class inbreeding of the major European powers.

Yet the first wave of colleges to be established were more the vestiges of an old colonial system than a new order. As Jerome Karabel's book *The Chosen* chronicles, they were a service for gentlemen, protective of class and sectarian distinctions, and, at first, were built to produce clergy from socially elite families. Harvard and Yale universities were born as tools of the Congregational Church; Columbia (originally known as King's College) served Episcopalians; Princeton did the same for Presbyterians; Rutgers was an affiliate of the Dutch Reformed Church; and Brown was for Baptists. Following the English model, each provided dormitories and dining halls, and enforced chapel attendance with a devotion to their particular evangelical doctrine.

There were differences in their respective markets for students, but social class was the most important factor for admissions throughout the nineteenth century. Harvard, modelled on Emmanuel College at the University of Cambridge, UK, required courses to be taught in Latin, the language of the church, well into the nineteenth century. It also ranked the social status of a student's family — a practice that continued into the twentieth century. Yale and other colleges did the same.

Women were generally not welcomed. When Harvard first offered instruction for women in 1894, it set up Radcliffe College, across the Charles River from Harvard, with largely its own faculty. Harvard itself did not invite women into its classrooms until the Second World War. As late as 1943, its governing

The upper crust: Harvard students in the 1870s were far from representative of the US public.

board refused to admit women to Harvard's medical school. Princeton and Yale, the two other members of the 'big three', stalled on coeducation at the undergraduate level until the 1960s.

Under its relatively progressive president Charles Eliot at the start of the twentieth century, Harvard, along with Yale and Princeton to a lesser degree, began to expand the scope of admissions — a conscious move to induce a more economically and culturally diverse student body that was more reflective of US society. The results were unimpressive. The big three formed the core, along with Columbia, of a cabal bent on marginally expanding access to certain preferred social and ethnic groups, while devising methods to exclude others.

One such tool was the College Entrance Examination Board (CEEB), which was established in the early 1900s. CEEB examinations, the forerunner of the SAT, became an important building block for a slight expansion of access for the middle class and, occasionally, the working class. It also offered a way of excluding 'unassimilable' populations — particularly Jews — that were unfamiliar with the cultural idioms built into the test.

Changes in the admission process in the 1920s helped to block the rest of the undesir-

ables. Abbott Lawrence Lowell, Eliot's successor at Harvard, pushed an admissions policy intended to serve the acceptable social elite — largely protestant, some Catholics, some German Jews, but definitely not the later wave of Russian Jews. Princeton and Yale were even more aggressive in maintaining their links to an acceptable social class.

The elite private institutions that dominated higher education along the northeastern seaboard didn't just ask prospective students to take standardized tests that were purposefully gauged to reflect the cultural norms of the well-bred Protestant. They also began to ask for pictures and non-academic information and, more overtly, placed a high value on accepting the male offspring of alumni.

Like Princeton and Yale, the cost of attending Harvard kept out many lower-income students. There were scholarships, but not many. Of the 3,500 students enrolled at Harvard in 1933, 84% were from extremely wealthy families. Those from the lower economic groups constituted perhaps less than 5% and were often academically high achievers of Jewish background.

Karabel's book provides a richly detailed version of this story, its subtitle promising the unveiling of a "hidden history" of admission

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

BETTMANN/CORBIS

and exclusion. Yet most observers of US higher education know of the elite nature and history of the big three and their role in educating the nation's rulers, past and present. The book is in part built on the shoulders of previous scholarly works, notably Marcia Graham Synnott's *The Half-Opened Door* (Greenwood, 1979) and Harold Wechsler's *The Qualified Student* (John Wiley, 1977).

What Karabel adds is an immersion into archival sources that allows him more fully to illuminate the voices of those who either set discriminatory admissions policies or struggled to change them. As Karabel observes, much has changed over the past forty or so years. The big three couldn't simply maintain their old allegiances and remain viable. Eventually, their leaders and influential alumni came to understand that greater inclusion meant they could play a larger role in society. They wanted their institutions to be more democratic and their students more academically talented.

In parallel with other universities, both public and private, Harvard, Princeton and Yale altered their admissions process to take more scholastically brilliant children from the middle and upper-middle classes. More important, they adopted affirmative-action policies to boost the number of minority students, initially focusing mainly on African-Americans.

There has, then, been a marked change for the better. The big three and a handful of other highly selective private institutions now educate a growing number of high-achieving minorities, some from lower-income backgrounds. Of the big three, Harvard has the highest percentage of undergraduates from ethnic minorities, about 28%. As in many other highly selective institutions, Asian-Americans are by far the largest minority group; African-Americans represent just 6.5%.

Some other things have not changed quite so much. Most minority students are from high-income families. Students from low-income families still go largely to public universities and colleges. Students from wealthy families still congregate at the most prestigious private institutions. Indeed, there is evidence that this trend is accelerating, reflecting, to some degree, the growing chasm between the rich and poor in the United States.

As Karabel notes, the big three are among the least diverse of the leading universities in terms of economic class. One reason is that admissions policies still favour the children of alumni. In 2002, 40% of such 'legacies' who applied to Harvard were admitted, compared with just 11% of all other applicants — a "birthright out of eighteenth-century British aristocracy, not twenty-first-century American democracy", one critic complained.

A limited supply of high-quality, prestigious, selective and increasingly wealthy private institutions, accompanied by growing demand both domestically and internationally, means the big three and their brethren will remain

élite and powerful. What is largely missing in Karabel's and other examinations of the big three is a parallel story: the rise of the public university movement in the United States and its huge impact on socio-economic mobility. The scale of that enterprise dwarfs that of the big three and other private universities. Arguably, the viability and fate of public uni-

versities will have a greater effect on the nation's democratic experiment and global competitiveness. But the star power and academic achievements of the big three continue to draw the most attention, obscuring this reality. ■

John Aubrey Douglass is at the Center for Studies in Higher Education, University of California, Berkeley, California 94720-4650, USA.

Cultural reflections

Hubris and Hybrids: A Cultural History of Technology and Science

by Mikael Hård & Andrew Jamison
Routledge: 2005. 335 pp. \$90 (hbk);
\$29.95 (pbk)

Howard P. Segal

It is a truism that culture, broadly defined, shapes science and technology as much as they shape culture. This once controversial position became the conventional wisdom decades ago, after purely internal histories of science and technology, followed by largely uncritical interpretations of their developments, were displaced as the dominant models.

In their excellent book *Hubris and Hybrids*, historians Mikael Hård and Andrew Jamison engage in a cultural assessment of science and technology. They replace the traditional 'heroic tale' of scientific genius with stories of the frequently mixed blessings of science and technology.

The 'hubris' of the title is reflected in James Watson's book *The Double Helix* (Atheneum, 1968), which recounts the race to discover the structure of DNA. In Watson's book the professional and monetary rewards were seen virtually as ends in themselves; there was a role for intuition along with conventional scientific methods; there was questionable treatment of peers; and there was little concern for the social and moral consequences of research.

For Watson, limits to either human intelligence or human power over nature had virtually disappeared. Yet Watson never denied his own flaws, and so helped to push scientific heroes off their traditional pedestals.

But even this account is too 'romantic' for Hård and Jamison, who seek even franker explorations of science from inside the laboratory — but only if paired with external (yet no less frank) analyses, such as Vandana Shiva's *Stolen Harvest* (South End, 2000). 'Hybrids' is the implicit theme of Shiva's book, which describes the way large corporations use the biotechnology derived from the genetic code. Some of these enterprises make huge profits while exploiting poor farmers, harming the environment, and undermining traditional balances between mankind and nature.

Hård and Jamison describe this story as a "tragedy" but wisely go beyond merely stressing the victimization. They never reduce their stories to wholesale good versus evil. Instead they focus on the growing convergence between science and technology into 'technoscience'. This is not simply about the elimination of most of the remaining barriers between scientific discovery and technological applications. It is also the story of changing meanings of being human, as we incorporate ever more technology within ourselves and our immediate surroundings. The authors discuss the possible cloning of people in the future, as well

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Frankenstein (Kenneth Branagh, left) refused to take responsibility for his creation (Robert De Niro).

as current issues such as the implantation of mechanical devices, the increased consumption of genetically engineered foods, a growing reliance on mobile phones and the Internet for daily communication, and endless modifications of the natural environment.

The authors invoke the influential concept of 'cyborgs': beings that are like humans in their ability to learn, feel and experience consciousness, but also like machines in having been 'programmed' to learn, feel and experience the world in only particular forms. Hence the authors' proper use of 'hybrids', a term they creatively apply to various contexts.

Hård and Jamison also provide useful summaries of the writings of earlier scholars, including Lewis Mumford, Siegfried Giedion, Lynn White and Raymond Williams, who all provided ground-breaking studies of science and technology in broad historical and cultural contexts, and Thomas Kuhn and Michel Foucault, who offered penetrating critiques of science and technology as being to varying degrees socially constructed. Hård and Jamison revisit, update and sometimes revise these earlier studies. By contrast, they criticize the founding editors of the leading journals of the history of science (George Sarton) and of the history of technology (Melvin Kranzberg) for promoting traditional uncritical views. Sarton's journal *Isis* may once have been guilty as charged, but Kranzberg's *Technology and Culture* was never so one-sided.

Far from being a critique of the excesses of only modern science and technology, *Hubris and Hybrids* is an extremely wide-ranging historical survey. Its coverage begins with the Scientific Revolution, Britain's Industrial Revolution, and the Enlightenment. More modern topics include technocracy movements, artistic uses of science and technology from William Morris to the film *The Matrix*, appropriate technology, the greening of corporate America and Europe, film and industrial design, and Asian developments. The richness of the authors' observations on these historical phenomena is exemplified by their comments on the medieval period: "eyeglasses and mirrors created opportunities to experience a technically mediated reality".

The authors hardly claim expertise in every area they discuss, but even so I was disappointed by their simplified comments on Mary Shelley's *Frankenstein*. Not only do Hård and Jamison follow most other commentators in wrongly describing Victor Frankenstein's unnamed and quickly abandoned creature as a "monster", but they also follow the crowd in wrongly characterizing Victor as a "mad scientist". Except in appearance, the "creature" — as he is usually called until the novel's later stages — is repeatedly portrayed as more human and humane than his creator. In my view, this should have been connected with the authors' own emphasis on humanity's changing identities. As for Victor, he is quite sane but is extraordinarily self-centred, as indifferent towards

his family and friends as he is to his creature. Ironically, his creature embodies Victor's missing moral compass.

Were Victor truly mad, he might well have escaped Shelley's actual target: his refusal to take responsibility for his research project. Here the authors missed the opportunity to use Frankenstein to bolster their own case. Neither work is a Luddite tract. Shelley argues that only if scientific experiments prove harmful to society should they be stopped. *Hubris and Hybrids* extends this same position

to inventors and engineers.

Recognizing that the relationship between the past and the future is different for historians from that for scientists, inventors and engineers, Hård and Jamison wisely offer no simple historical lessons, much less any silly predictions. What they provide instead are provocative and perceptive reflections that deserve to reach a wide general audience. ■

Howard P. Segal is in the Department of History, University of Maine, Orono, Maine 04469-5774, USA.

An inside view of the Universe

Theaters of Time and Space: American Planetaria, 1930-1970

by Jordan D. Marché II

Rutgers University Press: 2005. 266 pp. \$49.95

Shawn Cruzen

The spectacle of the night sky stretched overhead is the most breathtaking of natural wonders. Throughout time, people from every culture and from every part of the globe have experienced a sense of the infinite when confronted with the canopy of the Universe in a genuinely dark sky. However, the progression towards a more urbanized, industrialized society has brought with it a shroud of light pollution that has hidden this view from roughly half of Earth's population. But many city dwellers, especially in the United States, can still enjoy the splendour of a starry sky — by visiting their local planetarium, an island of sparkling celestial clarity (albeit artificial) in a sea of perpetual urban twilight.

In *Theaters of Time and Space*, author, science historian and planetarium veteran Jordan D. Marché II explores the evolution of planetaria from their inception in Germany to their proliferation across the United States. This account is both meticulous and colourful, and is sure to be enjoyed by anyone who is interested in astronomy, loves mechanical devices or has simply found inspiration under a planetarium's virtual starlight. The book outlines many of the social and cultural influences that fostered the spread of planetaria and their growth in popularity.

The concept of the planetarium was born from a confluence of ideas and technologies, including two early mechanical models of the Universe. The first of these, the orrery, uses a system of gears and wheels to demonstrate the motions of the Sun, Moon and planets. The second was a hollow rotating globe, large

enough to hold a small audience, with the stars and constellations painted on the interior to demonstrate celestial motions. With impetus from Oskar von Miller of the Deutsches Museum in Munich, an engineer named Walther Bauersfeld of the venerable Carl Zeiss optical company hit upon the idea of using projected images to show the motions of bodies in the Solar System against a fixed dome of painted stars. His colleague Werner Straubel then suggested optically projecting the stars as well. This engineering epiphany led to the genesis of the modern projection planetarium.

The Carl Zeiss company made the first planetarium for the Deutsches Museum on the roof of its own factory in Jena, Germany.



Star attraction: Zeiss projectors in the Adler Planetarium have given the Chicago public a glimpse of the heavens since 1930.

It opened in August 1923, and its abilities were demonstrated in the factory for a year before it was installed in the Munich museum the following August.

Marché chronicles the arrival of Zeiss planetaria in five major US cities between 1930 and 1939. Chicago's Adler Planetarium, featuring the Zeiss Model II projector, was the first, but others soon opened to enthusiastic audiences in Philadelphia, Los Angeles, New York and Pittsburgh.

Although Zeiss dominated early on, several imaginative inventors soon offered creative

ADLER PLANETARIUM

alternatives to the complex and expensive Zeiss systems. An interesting aspect of this story is the reaction of the directors of Zeiss planetaria to the arrival of competing systems. The book details how these influential individuals were able to stifle the propagation of less expensive planetaria until after the Second World War. It was then that Armand Spitz created and marketed a projector system that was smaller, simpler and more affordable than the Zeiss system. Spitz's legacy was to make planetaria much more numerous and hence more accessible to the public. Although attacked by Zeiss purists, the Spitz system was aided by a lack of direct competition and the dawning

of the space age. Spitz went on to become the world's largest producer of planetaria.

But *Theaters of Time and Space* is about more than just machinery. Marché touches on the human issues behind the birth of this industry. The book describes the emergence of the planetarium professional, a discipline containing elements of scientist, technician, teacher and entertainer. The history of women in the field is also examined, revealing the early difficulties they had breaking into this male-dominated profession. The book follows the careers of pioneering female planetarium directors and illustrates the growth of opportunities for women that came with the success

of the Spitz system. Marché also describes how the spiritual nature of astronomy inspired early philanthropic sponsorship of planetaria.

The book is a well written, thorough and enjoyable tribute to planetaria. It demonstrates their importance in encouraging interest in space science, providing communication between astronomers and the public, and promoting scientific literacy. It would be interesting to know how many astronomers can trace the inspiration that sparked their career to a planetarium visit.

Shawn Cruzen is director of the Coca-Cola Space Science Center, Columbus State University, Columbus, Georgia 31901, USA.

Dying for a drink

Evolution goes backwards in the latest Guinness advertisement.

Martin Kemp

Three young men in a bar are enjoying pints of Guinness, the dark Irish stout. Without warning, time reverses as they go on a dramatic journey back through evolutionary history. They are metamorphosed at dizzying pace into Edwardian gents who successively become Saxon, bronze-age cavemen, and, with a brief freeze in the Ice Age, apemen and apes. Then they become flying squirrels, terrestrial and aquatic mammals, fish, flightless birds, diminutive dinosaurs, and the mudskipper-like amphibians shown here — not exactly the correct ancestral line but it makes the point. During this helter skelter, Sammy Davis Jr intones the song *Rhythm of Life* from the musical *Sweet Charity*.

Finally, the mudskippers drink from their stagnant pool and one of them emits a disgusted burp. "Good things come to those who wait," we are told, as three pints of the dark stuff loom up, the central one labelled 'Guinness'. This latest arty offering from Guinness will saturate TV screens and cinema advertising, in Britain at least, for months to come.

Readers of Richard Dawkins will be reminded of *The Ancestor's Tale* (Weidenfeld Nicolson, 2004), in which the evolutionary story is told in reverse. The advert is of course linear, whereas Dawkins had complex trees to head each of his chapters, but the dust-jacket of the hardback edition parades just such a sequence.

Guinness advertising has for a long time been self-consciously promoted as an art form. The famous toucan with a pint perched precariously on its beak first appeared in



1935, designed by John Gilroy with copy written by Dorothy L. Sayers, noted scholar and author of detective tales: "Just think what Toucan do!" The punchline of the current ad, "Good things come to those who wait", is a knowing revival of an earlier slogan used in several campaigns. The company's website, www.guinness.com, makes no bones about exploiting the reputation of its famous ads, which have included the relatively recent "Pure genius" series.

The present dash through 3 billion years in a minute, and through a reputed million pounds of Guinness's advertising budget, is the responsibility of the ad agency AMV BBDO. Framestore created the dynamic visuals, under the direction of Daniel Kleinman, whose credits include James Bond movie titles.

The seamless morphing of creatures and settings involves an astonishing variety of techniques, ranging from compounds of dough and breakfast cereals cooked up at home to the most sophisticated laboratory programmes for three-dimensional animations. Location filming included a trip

to Iceland. The levels of visual consistency and conviction are startlingly high, whatever the nature of the original source material. Space, colour, light and shade, texture and motion conspire to blend the real and the artificial inseparably. For instance, real mudskippers were filmed in the studio sipping their muddy water. They were subsequently endowed with back fins and reanimated in a way that is not apparent to anyone who is not closely acquainted with the creatures.

The tone is humorous and ironic, both visually and in the implication that our ancestors were glumly waiting over the long years for the advent of a good pint. But viewers with different frameworks of belief are reacting in very different ways. For those educated to accept darwinian evolution as a fact of life, the whole sequence is seen as a virtuoso and memorable jeu d'esprit, centring on a tongue-in-cheek message about the true goal of natural selection.

For those who not only reject the validity of evolution and object to it being taught as a proven theory, appreciation of the wit will be obscured by hostility. The narrative will be seen as bowing to the scientific conspiracy to take godless evolution as fact. It isn't difficult to see how it could be criticized by those who are fuelling the current climate of anti-darwinism.

For a product marketed internationally, anything that grates on local sensitivities must be a matter of concern. All this reminds us that our own biological assumptions may be someone else's *bêtes noires*.

Martin Kemp is professor of the history of art at the University of Oxford, Oxford OX1 1PT, UK.

AMV BBDO

NEWS & VIEWS

OCEANOGRAPHY

The Atlantic heat conveyor slows

Detlef Quadfasel

Computer simulations predict that global warming will weaken the ocean circulation that transports heat from the tropics to higher latitudes in the North Atlantic. Such an effect has now been detected.

The Sun heats the tropics much more than the polar regions, but the resulting extremes of temperature are moderated by compensating heat circulation in the atmosphere and the ocean. Most notably, warm upper waters intrude far into the northern North Atlantic and contribute to the mild European climate. As Bryden *et al.* report on page 655 of this issue¹, that circulation has weakened over the past 50 years, providing worrying support for computer models that predict just such an effect in a world made warmer by greenhouse-gas emissions.

During their journey, the warm waters originating in the tropics release heat to the atmosphere and in consequence become denser. So they sink, and eventually return southwards at depth. This vertical 'overturning circulation' is driven by the north-south contrast in density and is thus vulnerable to processes affecting the densities at either end of the route. Besides changes in temperature, changes in salinity affect water density and so can alter the circulation. One consequence of global warming will be to inject more fresh water into the polar and sub-polar Atlantic², through enhanced precipitation, river run-off and melting of the Greenland ice cap. Increased input of fresh water reduces seawater density at high latitudes. Global atmosphere-ocean circulation models predict a slowdown of the ocean circulation in such circumstances³, with a consequent drop in temperatures over Europe of as much as 4 °C.

Bryden *et al.*¹ have analysed ocean temperature and salinity data, collected at intervals over the past five decades along a section at 25° N across the subtropical Atlantic. Their analysis provides the first observational evidence that such a decrease of the oceanic overturning circulation is well underway. Their approach is straightforward: the northward transport of water in the Gulf Stream system (the part concerned here being the Florida Current between the Gulf of Mexico and the Bahamas) has to be compensated by a return flow. This return flow consists of a wind-driven horizontal cell, which circulates clockwise in the subtropics, and the deep, southerly flow of the vertical overturning circulation (Fig. 1).

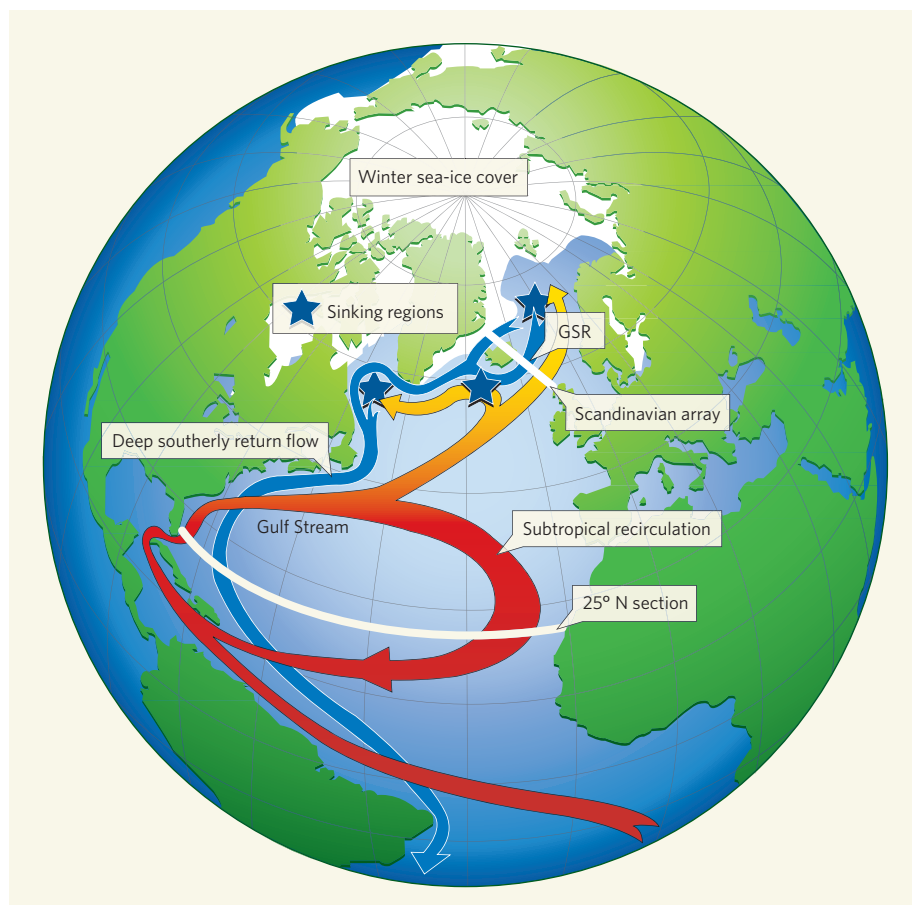


Figure 1 | The North Atlantic heat conveyor. Most warm waters in the upper ocean circulate clockwise in a giant horizontal swirl in the subtropics, but some flow farther north and cross the Greenland-Scotland Ridge (GSR). This branch warms the northern North Atlantic and Europe, and keeps most of the Nordic Seas free of ice. Here the water sinks (indicated by the star) and flows back southwards at depth, mostly down the western edge of the Atlantic basin. The Scandinavian monitoring array tracks only the northern limb of the overturning circulation, but more deep water is added south of the GSR and in the Labrador Sea (stars). The 25° N section covers all of the overturning circulation, and also includes the horizontal recirculation in the subtropics. According to Bryden and colleagues' results¹, the former is weakening and the latter strengthening.

Because of the large temperature difference between its upper and lower branches, the overturning circulation contributes about 90% to the oceanic south-north heat fluxes in the North Atlantic and is thus most significant for European climate and its variability.

Bryden and colleagues estimate the strength of these two circulation cells from

repeated observations of the oceanic density field. They apply the so-called thermal wind relation, in which the internal pressure field is calculated from the density distribution and balanced with the Coriolis force that arises when a fluid is moving on the rotating planet. The result is alarming: a significant shift from the vertical to the horizontal circulation cell

has occurred, with the vertical cell dominating until the early 1990s and the horizontal cell thereafter. Altogether, the overturning circulation has decreased by more than 30%, with the largest effect seen in the lower part of the deep water.

But how solid are these results? The findings are based on just five snapshots of the circulation, taken in 1957, 1981, 1992, 1998 and 2004, and along one latitudinal section. Higher-frequency variability (such as eddies or waves), at the ends of the section at the African coast and the Bahamas, may obscure the detection of long-term change. And the uncertainty of the estimates given is high, so the magnitude of the decline may well be smaller than suggested by the calculations. Against that, the declining trend itself is statistically significant. Also, the observed density structure of the deep waters has changed; this structure is not affected by short-term variability, and so supports Bryden and colleagues' conclusions.

Further support comes from other observations. Based on direct measurements of water flux, along with model calculations, Hansen *et al.*⁴ have reported a 20% reduction in the overflow from the Nordic Seas across the Greenland–Scotland Ridge into the deep North Atlantic over the past 50 years. These overflows contribute about one-third to the overturning circulation and feed the densest part of the vertical cell — the same part in which the largest reduction of water transport was observed at 25° N. At the same time, the overflow waters and in turn the deep waters of the North Atlantic have significantly freshened⁵, thereby reducing the large-scale density gradient driving the overturning circulation.

The implications of these observations are considerable. Palaeoclimate records show that northern air temperatures can drop by up to 10 °C within decades⁶, and that these abrupt changes are intimately linked to switches in the ocean circulation⁷. Increased freshwater input into the Nordic Seas will initially weaken the circulation only slowly. But when a certain threshold is reached, the circulation may jump abruptly to a new state in which there is little or no heat flux to the north. The system is highly nonlinear and will not immediately switch back to the warm mode when the freshwater input weakens.

Although most model studies agree on the way the overturning circulation will develop under global-warming conditions³, direct observations of the relevant fluxes of water volume and heat are still sparse. Scandinavian oceanographers are monitoring the northern limb of the vertical cell⁸, and a basin-wide array of moored temperature and salinity recorders has recently been deployed along the 25° N section⁹. Such observations, combined with modelling activities, are crucial for providing an early warning of a possible breakdown of the overturning circulation — which, if it occurs, would have devastating

effects on socio-economic conditions in the countries bordering the eastern North Atlantic.

Detlef Quadfasel is at the Zentrum für Meeres- und Klimaforschung, Institut für Meereskunde, Universität Hamburg, Bundesstrasse 53, D-20146 Hamburg, Germany.
e-mail: quadfasel@ifm.zmaw.de

1. Bryden, H. L., Longworth, H. R. & Cunningham, S. A. *Nature* **438**, 655–657 (2005).

2. Curry, R., Dickson, B. & Yashayaev, I. *Nature* **426**, 826–829 (2003).
3. Intergovernmental Panel on Climate Change *Climate Change 2001: The Scientific Basis* (eds Houghton, J. T. *et al.*) (Cambridge Univ. Press, 2001).
4. Hansen, B., Turrell, W. R. & Østerhus, S. *Nature* **411**, 927–930 (2001).
5. Dickson, R. *et al.* *Nature* **416**, 832–837 (2002).
6. Dansgaard, W. *et al.* *Nature* **364**, 218–220 (1993).
7. Ganapolski, A. & Rahmstorf, S. *Nature* **409**, 153–158 (2001).
8. Østerhus, S., Turrell, W. R., Jónsson, S. & Hansen, B. *Geophys. Res. Lett.* **32**, L07603; doi:10.1029/2004GL022188 (2005).
9. Srokosz, M. A. *EOS Trans. AGU* **85**, 78, 83 (2004).

BIOPHYSICS

Assembly line inspection

Sarah A. Woodson

Many of the biochemical events that occur in a cell are performed by huge complexes of proteins and nucleic acids. A cunning approach promises to show how the components convene to make a functioning 'machine'.

The cell's macromolecular machines contain dozens or even hundreds of components. But unlike man-made machines, which are built on assembly lines, these cellular machines assemble spontaneously from their protein and nucleic-acid components. It is as though cars could be manufactured by merely tumbling their parts onto the factory floor. Knowing how cellular complexes organize themselves is crucial for understanding molecular evolution and for engineering materials that can mimic their properties. In this issue, Talkington and colleagues (page 628)¹ use isotopically labelled proteins and mass spectrometry to follow the assembly of one macromolecular complex — the 30S ribosome — in real time. The elegance of their experimental design should allow it to be adapted to a wide range of such complexes, offering a new approach to the study of cellular dynamics.

Ribosome assembly is fascinating, because its components must form a stable yet flexible platform for protein synthesis. The small (30S) subunit of the 70S ribosome found in bacterial cells consists of the 16S rRNA (1,542 nucleotides) and 20 unique proteins, which interact through highly specific interfaces (reviewed in ref. 2). Actively growing cells demand many thousands of ribosomes, whose synthesis consumes a large fraction of the cell's metabolic energy^{3,4}. So ribosome assembly must be efficient as well as precise.

More than 30 years ago, Nomura and his colleagues⁵ demonstrated that 30S ribosomal subunits could be made *in vitro* from purified proteins and 16S rRNA. By varying the order in which the proteins were mixed with the rRNA, a hierarchy of protein–RNA interactions was defined. This assembly 'map' explained the remarkable specificity of ribosome reconstitution, because only a few

proteins could bind to the naked 16S rRNA and initiate assembly. The binding of these primary initiating proteins then enable the next ones in the hierarchy to bind, and so on. Biochemical and crystallographic experiments showed that this cooperativity of assembly derives mostly from protein-induced changes in the structure of the 16S rRNA, rather than from direct protein–protein contacts (reviewed in ref. 6).

But to fully understand the dynamics of a macromolecular complex, one must follow its components in real time. Until now, the kinetics of 30S ribosome assembly has been studied by monitoring changes in the conformation of the 16S rRNA^{7,8}. Talkington *et al.* took a completely different approach, measuring the rate at which each protein is added to the 16S rRNA *in vitro*. They used a 'pulse–chase' strategy to introduce isotopically labelled proteins within a certain time window (Fig. 1). The 16S rRNA was incubated with a mixture of ribosomal proteins isolated from cells grown in ¹⁵N-containing media, and then excess ¹⁴N-labelled proteins were added (the 'chase'). Fully reconstituted 30S particles were purified to remove any nonspecifically bound proteins. Next, the ratio of ¹⁵N to ¹⁴N protein was measured by subjecting the entire mixture to MALDI–TOF mass spectrometry. Not only were most of the 20 proteins identified in the mass spectrum, but for each protein the ratio of ¹⁵N- to ¹⁴N-labelled polypeptide was accurately determined. By varying the length of the ¹⁵N pulse, the authors obtained rate constants for the association of 17 of the 20 proteins with 16S RNA.

This method of pulse–chase quantitative mass spectrometry (PC/QMS) cleverly sidesteps some of the technical problems that can plague studies of large macromolecular

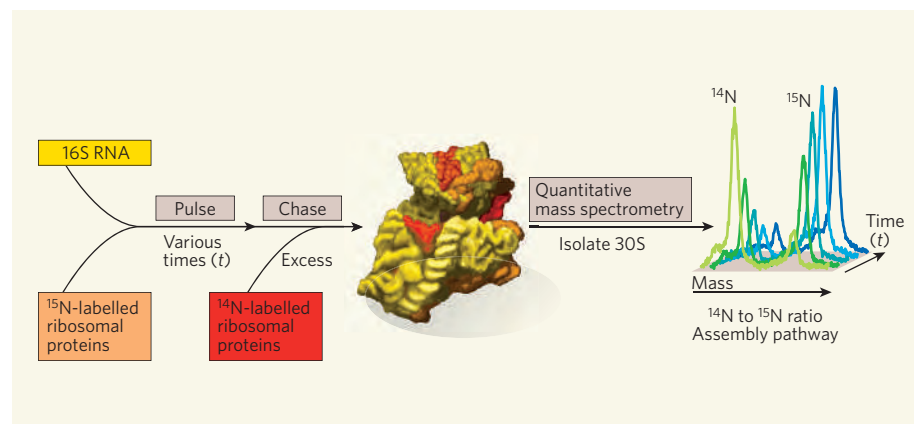


Figure 1 | Tracking assembly of the 30S ribosome in real time. 16S rRNA and twenty ^{15}N -labelled 30S proteins are mixed together *in vitro* to begin assembly of the ribosome. After a set period of time (t), further binding of ^{15}N -protein is inhibited by the addition of excess ^{14}N -labelled proteins. Once the assembly reaction is complete, the 30S complexes are isolated by centrifugation and analysed by mass spectrometry. The isotopic ratios of the ribosomal proteins at various times determine the order in which they join the complex, giving the 30S ribosome assembly map.

complexes. First, by purifying 30S particles before mass spectrometry, contributions from 'dead-end' intermediates are excluded. Second, simultaneous analysis of all 20 small-subunit proteins allowed the authors to use native proteins from cells grown in ^{15}N media, avoiding the need to overexpress individual proteins. The PC/QMS approach is attractively simple, and in principle can probably be adapted to examine many large RNA and protein complexes. The main requirements are reconstitution of the functional complex and the ability to label proteins using heavy isotopes or some other mass tag.

Given what we know about 30S ribosome assembly, what can be learnt from the protein association rates? Nomura and colleagues⁹ observed that reconstitution at 30 °C stalls at an intermediate (RI) that lacks a subset of the 30S proteins. Thermal activation of RI at 42 °C has been correlated with refolding of the 16S rRNA near the centre of the 30S subunit⁸. If the rate of 30S assembly is limited by a single step, proteins that bind to the 16S rRNA after this rate-determining step should associate more slowly than those that bind before it and show a similar activation enthalpy.

By contrast, Talkington *et al.* find that neither the apparent association rate nor the activation enthalpy of binding correlate with the position of the protein in the assembly map. Instead, their results suggest that multiple rate-limiting steps control the kinetics of ribosome assembly. This is consistent with individual ribosomes taking alternative routes to the final structure.

Further work is needed to determine whether 30S assembly requires specific intermediates, or whether reconstitution intermediates represent meta-stable structures in a complex free-energy landscape. In simpler RNAs, meta-stable folding intermediates are observed when the RNA is kinetically trapped in misfolded structures^{10,11}. However, some RNA molecules fold by

alternative paths that avoid these trapped intermediates^{10,12}. In the PC/QMS experiments, such stalled intermediates might have been missed because ^{15}N proteins that bind weakly to stalled intermediates could be washed out by the ^{14}N chase.

In the soft world of biological materials, cooperativity and specificity are achieved by the induced fit of molecular interfaces; that is, as two or more components come into contact they mould around one another to create stronger, more specific junctions. The idea that ribosome assembly can follow more than one path is consistent with redundant cooperative linkages in the assembly map⁵. These cooperative linkages ensure that individual complexes are assembled completely. They also create alternative kinetic paths that make the assembly process itself more robust. In the ribosome, these interactions have been fine-tuned through billions of years of evolution, providing a clear window into the world of cellular machines.

Sarah A. Woodson is in the T. C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA.
e-mail: swoodson@jhu.edu

1. Talkington, M. W. T., Siuzdak, G. & Williamson, J. R. *Nature* **438**, 628–632 (2005).
2. Ramakrishnan, V. & Moore, P. B. *Curr. Opin. Struct. Biol.* **11**, 144–154 (2001).
3. Nierhaus, K. H. *Biochimie* **73**, 739–755 (1991).
4. Warner, J. R. *Trends Biochem. Sci.* **24**, 437–440 (1999).
5. Nomura, M. *Science* **179**, 864–873 (1973).
6. Culver, G. M. *Biopolymers* **68**, 234–249 (2003).
7. Powers, T., Daubresse, G. & Noller, H. F. J. *Mol. Biol.* **232**, 362–374 (1993).
8. Holmes, K. L. & Culver, G. M. *Nature Struct. Mol. Biol.* **11**, 179–186 (2004).
9. Traub, P. & Nomura, M. *J. Mol. Biol.* **40**, 391–413 (1969).
10. Pan, J., Thirumalai, D. & Woodson, S. A. *J. Mol. Biol.* **273**, 7–13 (1997).
11. Treiber, D. K. & Williamson, J. R. *Curr. Opin. Struct. Biol.* **9**, 339–345 (1999).
12. Zhuang, X. *et al. Science* **288**, 2048–2051 (2000).



50 YEARS AGO

This is the centenary year of the Meteorological Office, and through its long record meteorologists have been confronted with dynamical problems of such complexity that their solution has sometimes seemed beyond hope. In recent years, however, nothing less than a revolution has taken place...

The staff of the Napier Shaw Research Laboratory of the Meteorological Office...have shown that weather maps giving isobars and upper air contours may be predicted for 24 hr. ahead.

From *Nature* 3 December 1955.

100 YEARS AGO

Nature Through Microscope and Camera. By Richard Kerr. One of the many ways of beginning the study of natural science is with "beauty-feast" — of flowers or birds, of shells or gems, of anything — for all natural things are beautiful, in their proper setting at least. It is an old-fashioned mode of approach, commending itself to children and simple minds, but one which often leads far beyond aesthetic pleasure to the joy of understanding. It affords a dynamic to investigation, and fosters a healthy reverence for things. Indeed, if we had to choose, we should prefer admiration without science to science without admiration. But a simple book like that before us shows that there is no necessary antithesis; it is a disclosure of beautiful things, and yet within its limits it is quite scientific.

The author's aim is to illustrate with well chosen examples the beauty of minute structure, the beauty which the microscope discloses, and he is to be congratulated on his success... we are here brought into close quarters with the familiar, with diatoms and Foraminifera, the whelk's radula and the barnacle's cirri, the butterfly's "tongue" and the scales of the sole... The photographs were taken by Mr. Arthur E. Smith, and are certainly among the finest that have ever been published.

From *Nature* 30 November 1905.

50 & 100 YEARS AGO

CELL BIOLOGY

A greasy grip

Anthony G. Lee

How do the lipids and proteins of the cell membrane interact to create a functioning barrier for the cell?
A high-resolution structure of a membrane protein reveals intimate contacts with its lipid neighbours.

The design of a biological membrane is beautifully simple: a lipid bilayer provides the basic barrier, and into this are plugged a variety of membrane proteins. Each protein is designed to carry out some particular function for the cell — moving specific molecules in or out of the cell, say, or sensing the environment. The trick is to make sure that the lipid bilayer and the membrane proteins are mutually compatible: the proteins must be able to operate well in the environment provided by the lipid bilayer, and insertion of the proteins into the bilayer must not make it leaky, or the permeability barrier would be destroyed. In other words, the key to an effective membrane is to get the packing of the lipids and proteins right. On page 633 of this issue, Walz and colleagues¹ present high-resolution structural data that clearly illustrate how this packing is achieved.

Thirty years ago, studies of biological membranes using electron spin resonance detected a population of lipid molecules whose fatty acyl chains were conformationally disordered². These chains were thought to belong to the lipid molecules — called boundary or annular lipids — that contacted membrane proteins. But not everyone was convinced, hence the significance of Walz and colleagues' work; nothing is as compelling as actually seeing something with your own eyes.

The authors have used electron crystallography to determine the structure of a protein called lens-specific aquaporin-0 (AQP0) when it is immersed in an artificial lipid bilayer¹. AQP0 is the most abundant protein in the plasma membranes of the fibre cells that make up the bulk of the lens in the human eye. Functionally, AQP0 mediates rapid movement of water into and out of the fibre cells, but it also has a structural role, forming membrane junctions between fibre cells. Under the right conditions (when at least a fraction of the AQP0 molecules are partially proteolytically cleaved) the AQP0–lipid-bilayer system forms two-dimensional crystals consisting of a pair of closely spaced membranes.

The structure that Walz and colleagues obtained from these crystals is a remarkable

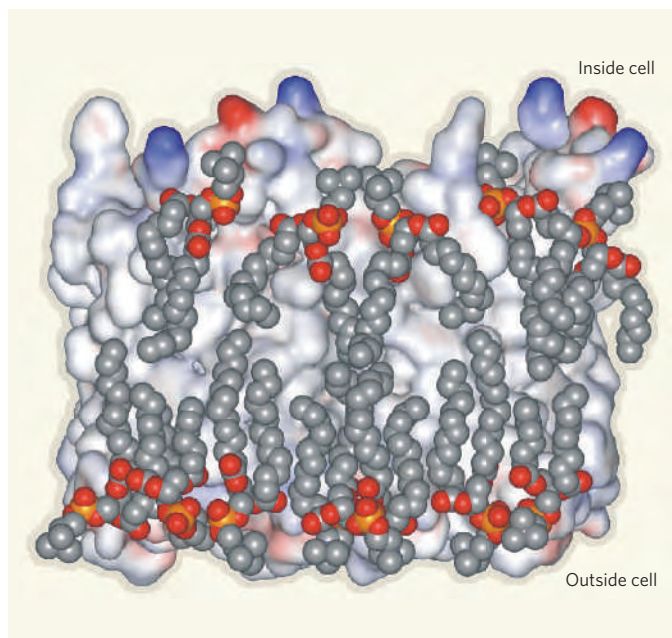


Figure 1 | Annular lipids. A side view of the electron crystallographic structure reported by Walz and colleagues¹ showing one face of the AQP0 tetramer with lipid molecules forming a bilayer shell around the protein. The AQP0 tetramer is shown as a surface plot (the lighter background molecule), with red representing regions of negative charge, blue, regions of positive charge, and grey, uncharged regions. The lipid molecules are shown in space-fill format (the molecules in the foreground). The charged lipid headgroups (oxygen, red; phosphorus, orange) and the lipid fatty acyl chains form a bilayer with almost uniform thickness around the protein. Presumably, in the membrane, lipid fatty acyl chains will cover the whole of the hydrophobic surface of the protein; only the most ordered of the lipid molecules will be resolved in the crystallographic structure.

technical achievement: it is the highest-resolution structure ever produced from electron crystallography, and only the second to show the lipid environment of a membrane protein. The other structure showing this is of bacteriorhodopsin from the purple membranes of a photosynthetic bacterium³.

Walz and colleagues' structure shows the AQP0 molecules in their closed form, organized as complexes of four AQP0 units (tetramers) in each lipid bilayer. The flat extracellular surface of a tetramer in one bilayer contacts the extracellular surface of another AQP0 tetramer in the adjacent bilayer. This 'head-to-head' packing links the two adjacent bilayers to form a structure that looks very like an *in vivo* cell–cell junction⁴.

The artificial lipid bilayer is made of dimyristoylphosphatidylcholine, a molecule with a 'head' containing a negatively charged phosphate group and a positively charged choline group, and two 'tails' of fatty acyl

chains, each with 14 carbons. In Walz and colleagues' structure, the lipids can be seen forming a shell around each AQP0 tetramer (Fig. 1). The tetramers are separated by a shell of lipid molecules, just one molecule thick. This makes the lipid molecules unusual in that most are in contact with two protein molecules, one from each of the adjacent tetramers.

Perhaps this partly explains why so many of the lipid molecules are resolved in the structure. In the only other high-resolution structure showing a large number of lipid molecules, that of bacteriorhodopsin, lipid molecules similarly mediate ordered packing of the protein molecules, in this case of trimers of bacteriorhodopsin³. The crystalline array of closely packed AQP0 tetramers seen in these reconstituted systems is very like that seen in the native lens fibre membrane⁵.

The striking thing about the structure of AQP0 is the excellent packing of the lipid fatty acyl chains against the rough surface of the AQP0 molecule. For some lipid molecules, this is achieved with almost straight (all-*trans*) fatty acyl chains, but for others considerable distortion of the

chains is necessary for them to wrap around the bulky side-chains of the protein (Fig. 1). The lipid headgroup conformations also differ markedly between the various lipid molecules with some lipid headgroups being oriented almost parallel to the surface of the membrane, as occurs in crystals of phosphatidylcholines⁶, but with others being oriented almost vertically.

What the structure makes clear is that the AQP0 surface is not covered by a set of uniform binding sites for phospholipid molecules. The binding of lipids to AQP0 is therefore very unlike the interactions between many other phospholipids and proteins, where particular lipids bind to unique binding sites on the surface of the protein, a classic example being the highly specific binding of phosphatidylinositols to protein domains called PH domains. Rather, the picture presented is of a number of phospholipid molecules, each with their fatty acyl chains interacting with the

large, rough, hydrophobic surface of the protein; their charged headgroups interact with the charged residues flanking the hydrophobic surface. Another interesting feature of the structure is that the thickness of the lipid bilayer around the AQP0 tetramer is rather uniform, despite the fact that the structures adopted by the individual lipid molecules are very different. The lipid bilayer observed within the recently published structure of another membrane protein similarly seems to be of constant thickness⁷.

The picture given here is, of course, of a membrane frozen in time — the structure was determined at low temperature. At normal temperatures the lipid bilayer would be more fluid; lipid molecules would probably be rapidly entering and leaving the annular shells around the protein molecules, and the encounter between a particular lipid molecule and a particular protein molecule would be brief⁸. Nevertheless, the structures adopted by the lipid molecules when they are on the protein surface would be much like those pictured here, and fast swapping of lipid molecules between the annular shell and the bulk lipid bilayer would not change the environment 'experienced' by the protein. The protein would always experience lipid molecules in the disordered states shown in the structure of Walz and colleagues¹.

Crystalline membranes of AQP0 can be formed with a variety of lipids other than dimyristoylphosphatidylcholine¹. It might therefore be possible to use this system to answer several questions exercising the minds of membranologists. If AQP0 were reconstituted with a longer chain lipid, how would the long fatty acyl chains of the lipid molecules distort to ensure that the hydrophobic thickness of the lipid bilayer matched the hydrophobic thickness of the protein? How would lipids such as phosphatidylethanolamine that prefer to form non-bilayer structures interact with AQP0? Would changing the structure of the lipid result in any change in the structure of the AQP0 molecules? With answers to these and similar questions, we can start to understand how lipid and protein molecules coevolved to form membranes that are fit for their purpose.

Anthony G. Lee is in the School of Biological Sciences, University of Southampton, Southampton SO16 7PX, UK.
e-mail: a.g.lee@soton.ac.uk

1. Gonen, T. *et al.* *Nature* **438**, 633–638 (2005).
2. Jost, P. C., Griffith, O. H., Capaldi, R. A. & Vanderkooi, G. *Proc. Natl Acad. Sci. USA* **70**, 480–484 (1973).
3. Luecke, H., Schober, B., Richter, H. T., Carttailler, J. P. & Lanyi, J. K. *J. Mol. Biol.* **291**, 899–911 (1999).
4. Gonen, T., Sliz, P., Kistler, J., Cheng, Y. & Walz, T. *Nature* **429**, 193–197 (2004).
5. Zampighi, G., Simon, S. A., Robertson, J. D., McIntosh, T. J. & Costello, M. J. *J. Cell Biol.* **93**, 175–189 (1982).
6. Pascher, I., Lundmark, M., Nyholm, P. G. & Sundell, J. *Biochim. Biophys. Acta* **1113**, 339–373 (1992).
7. Murata, T., Yamato, I., Kakinuma, Y., Leslie, A. G. W. & Walker, J. E. *Science* **308**, 654–659 (2005).
8. Lee, A. G. *Biochim. Biophys. Acta* **1612**, 1–40 (2003).

PLANETARY SCIENCE

Clays in the history of Mars

Horton Newsom

The stream of revelations from Mars continues. The latest news — the discovery of clays in ancient terrains — helps to fill in the picture of the past existence of liquid water on the planet's surface.

Thanks to three orbiting spacecraft and two rovers that continue to return data, we now have a great deal of evidence that abundant surface water once existed on Mars. But until now something has been missing. Water would be expected to alter the surface minerals to form clays, yet that hallmark of aqueous action seemed to be absent.

The mystery has been solved by new results from the OMEGA near-infrared spectrometer experiment on the Mars Express spacecraft. Clays do indeed exist on Mars, but only in particular places. The OMEGA instrument's initial discovery¹ of clays on ancient martian surfaces, published earlier this year, is now followed by the more comprehensive account by Poulet *et al.* that appears on page 623 of this issue². Clays are a subgroup of the phyllosilicates, water-bearing silicate minerals with layered structures, and Poulet and colleagues' observations show that different types of phyllosilicates are present in some outcrops of the ancient highlands of Mars. In addition to clearing up the mystery of the missing clay, this observation provides evidence for a drastic change in the chemistry of surface processes early in martian history.

Apart from circumstantial evidence for the role of water in forming gulleys at high latitudes³, there are no other indications of large amounts of liquid water at the surface of Mars today. By contrast, there has been plenty of evidence for abundant past water on Mars, provided by images of geomorphic features that seem to have been produced by the action of

liquid water. Recent discoveries include deltas and previously unknown channel systems, suggesting that abundant surface run-off occurred in the distant past.

Indications of the previous existence of water on Mars have also come from a different source — data sent back by the Mars Exploration Rover Opportunity from the Meridiani Planum landing site. Most notably, the discovery⁴ of haematite and sulphate minerals such as jarosite on the surface at Meridiani suggests that liquid water must have been present at the surface and as ground water at some stage.

On Earth, the reaction of water, especially hot water, with rock leads to the formation of water-bearing alteration products, including clays. So clay minerals should be abundant on Mars, especially early in the planet's history when water was present and there was heat from volcanoes and large impact craters^{5–8}. On the basis of earlier spectra, a few voices have indeed argued that hydrated silicate clay minerals are present^{9,10}. The consensus, however, also taking into account the recent data from the Mars Exploration Rovers, has been that the mineralogy of the martian surface is characterized by acid-sulphate processes.

Compared with Earth, where a wide range of minerals is formed by aqueous action, this acid-sulphate environment limits the number of minerals that can be created. In particular, clays are not usually produced in such environments^{11,12}. It also seems that these acid-sulphate conditions have been prevalent for much of Mars' history, in that there is a lack of

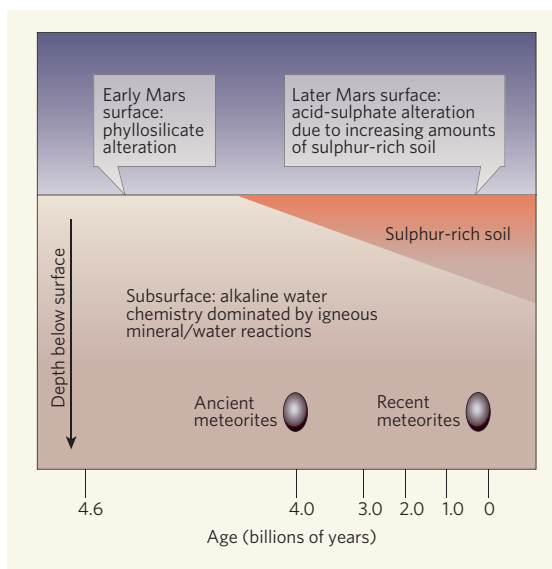


Figure 1 | Water-rock interaction on Mars. The data from Poulet *et al.*² suggest that phyllosilicates, water-bearing silicate minerals that include clays, may have formed early in Mars' history. The surface later became acidic, leading to the formation of sulphate-rich aqueous-alteration minerals. Martian meteorites that have reached Earth were ejected by impacts and are derived from below the planet's surface. The alteration phases in meteorites of both ancient and recent origin generally show evidence for modification in a more alkaline aqueous environment¹⁶. Presumably, then, the sulphur-rich acidic material that has accumulated on the surface is not present at depth.

abundant carbonate deposits on the planet's surface¹³. Carbonate deposits are expected to be present as the final stage of evaporation of large bodies of water, because these deposits are found on Earth in such circumstances¹⁴. On Mars, an ocean is thought to have covered the northern plains, but if so it has left no carbonate imprint.

Did surface environments ever exist on Mars where abundant clay minerals could form? They did — but several billion years ago. Poulet *et al.*² have identified the characteristic spectra of a range of phyllosilicates, including the iron-rich smectite nontronite, the aluminium-rich smectite montmorillonite, and an iron–magnesium chlorite mineral called chamosite. These alteration minerals were found in the cratered highlands of Mars, which formed very early in Mars' history, perhaps more than 3.5 billion years ago, based on the impact cratering record. So this suggests that in the distant past the martian surface was more like Earth's current state than the acidic, sulphur-rich environment that the Mars Exploration Rovers are documenting today.

The discoveries of phyllosilicates² and of sulphates⁴ are leading to a more detailed understanding of the interaction between water and surface materials on Mars. The presence of phyllosilicates in the ancient highlands suggests that Earth-like conditions existed well before 3.5 billion to 4 billion years ago. During later martian history, it seems that the surface became more acidic, suppressing the formation of phyllosilicates and carbonates, and leading to the haematite and sulphates spectacularly observed at Meridiani by Opportunity.

What about conditions below the martian surface? Meteorites originating from Mars can provide some clues. Most of those known formed comparatively recently (about 0.18 billion years ago) and they contain aqueous alteration products. These products include carbonates, which were produced from igneous rocks under alkaline conditions, rather than in an acidic environment in contact with the surface soil^{15,16}. Given that these meteorites are so young, they were probably altered at depth below the acidic surface layer.

A new view of martian chemistry is summarized in Figure 1. According to this view, there was a change in surface chemistry over time from a less acidic to a more acidic regime. And in the crust, at the depths from which the recent meteorites are probably derived, the aqueous chemistry continues to be dominated by alkaline chemical reactions between igneous minerals and water.

Among other consequences, the discovery of phyllosilicates may influence the search for past environments suitable for life, such as fossil hydrothermal systems. Clays may be associated with the origin of life itself, which perhaps started with the formation of complex organic molecules on a clay mineral substrate; their existence on Mars may therefore be a

consideration in site selection for future explorations such as the next rover mission, the Mars Science Laboratory, to be launched in 2009. In the meantime, we can look forward to yet more discoveries — from further near-infrared studies by both the current OMEGA spectrometer and the Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) instrument on the Mars Reconnaissance Orbiter, which goes into orbit around the planet in March 2006.

Horton Newsom is in the Institute of Meteoritics and the Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico 87131, USA.
e-mail: newsom@unm.edu

1. Bibring, J.-P. *et al. Science* **307**, 1576–1581 (2005).

2. Poulet, F. *et al. Nature* **438**, 623–627 (2005).
3. Malin, M. C. & Edgett, K. S. *Science* **288**, 2330–2335 (2000).
4. Squyres, S. W. *et al. Science* **306**, 1709–1714 (2004).
5. Newsom, H. E., Hagerty, J. J. & Thorsos, I. E. *Astrobiology* **1**, 71–88 (2001).
6. Gulick, V. C. *J. Geophys. Res.* **103**, 19365–19388 (1998).
7. Rathbun, J. A. & Squyres, S. W. *Icarus* **157**, 362–372 (2002).
8. Abramov, O. & Kring, D. A. *Lunar Planet. Sci. Conf. XXXVI*, abstr. 1048 (2005).
9. Bishop, J. L. *et al. Icarus* **117**, 101–119 (1995).
10. Wyatt, M. B. & McSween, H. Y. Jr *Nature* **421**, 712–713 (2003).
11. Morris, R. V., Graff, T. G., Mertzman, S. A., Lane, M. D. & Christensen, P. R. *Sixth Int. Conf. on Mars* abstr. 3211 (2003).
12. Morris, R. V. *et al. Science* **305**, 833–836 (2004).
13. Christensen, P. R. *et al. J. Geophys. Res.* **106**, 23823–23871 (2001).
14. Catling, D. C. *J. Geophys. Res.* **104**, 16453–16469 (1999).
15. Bridges, J. C. *et al. Space Sci. Rev.* **96**, 365–392 (2001).
16. Zolotov, M. Yu., Shock, E. L., Niles, P. & Leshin, L. *Second Conf. on Early Mars: Geologic, Hydrologic, and Climatic Evolution and the Implications for Life* abstr. 8036 (2004).

CELL BIOLOGY

Protein choreography

Mara C. Duncan and Gregory S. Payne

Just under the cell surface, proteins engage in an intricate ballet to drive a transport process called endocytosis. Much is known about the individual dancers, but now the choreography is revealed.

Endocytosis is the process by which cells gulp up small patches of their outer plasma membrane, sucking them inside to form small membrane-enclosed vesicles. These bubble-like structures provide a transport mechanism for carrying proteins that were embedded in the outer membrane, extracellular molecules associated with some of those proteins and small amounts of extracellular fluid¹. The process can affect both normal and disease states of the cell through its roles in nutrient acquisition, cell growth, neural transmission and the entry of viruses into cells.

One well-characterized endocytosis pathway involves the protein clathrin, which coats the vesicles. The individual activities and interactions of many of the protein factors in the clathrin-mediated pathway have been characterized, but how these myriad interactions and activities are integrated to drive endocytosis is not clear. Drubin and colleagues², writing in *Cell*, now provide a temporal and spatial framework for the protein dynamics responsible for endocytic vesicle formation in living yeast cells.

More than 40 years ago, electron microscopy provided the first glimpse of clathrin-coated vesicles and half-formed pockets in the plasma membrane³. These types of static image, combined with biochemical and molecular genetic analysis of clathrin and other coat components, engendered a generally accepted model of endocytic vesicle formation¹. The clathrin coats assemble as polygonal lattices on the inside of the plasma

membrane. Other coat components bind to the intracellular regions of certain plasma membrane proteins, capturing the proteins and ensuring that they will become cargo in the resulting vesicle. The assembly of the coat components is also associated with inward protrusion of the plasma membrane to generate a coated invagination. Finally, the invaginated membrane pinches off to generate a free, coated vesicle, which then sheds its coat as it moves farther into the cell.

Drubin and colleagues² have examined endocytosis in the yeast *Saccharomyces cerevisiae* by tagging proteins implicated in the process with green fluorescent protein and following their movements in living cells using fluorescence microscopy⁴. This team previously demonstrated sequential assembly and disassembly of several proteins at sites of endocytosis⁵. By monitoring positions of assembled proteins at different times, the group defined three stages in vesicle formation: an initial phase, in which a patch of coat proteins forms but does not move; a second phase, when the endocytic structure moves relatively slowly to a short distance away from the plasma membrane; and a final phase, in which the structure moves rapidly into the cell interior. These phases were proposed to reflect coat assembly, coated membrane invagination and release of the free vesicle.

What emerges from the more comprehensive analysis in the current study is a detailed view of the dynamics of vesicle formation, governed by four sets of proteins, or 'modules'

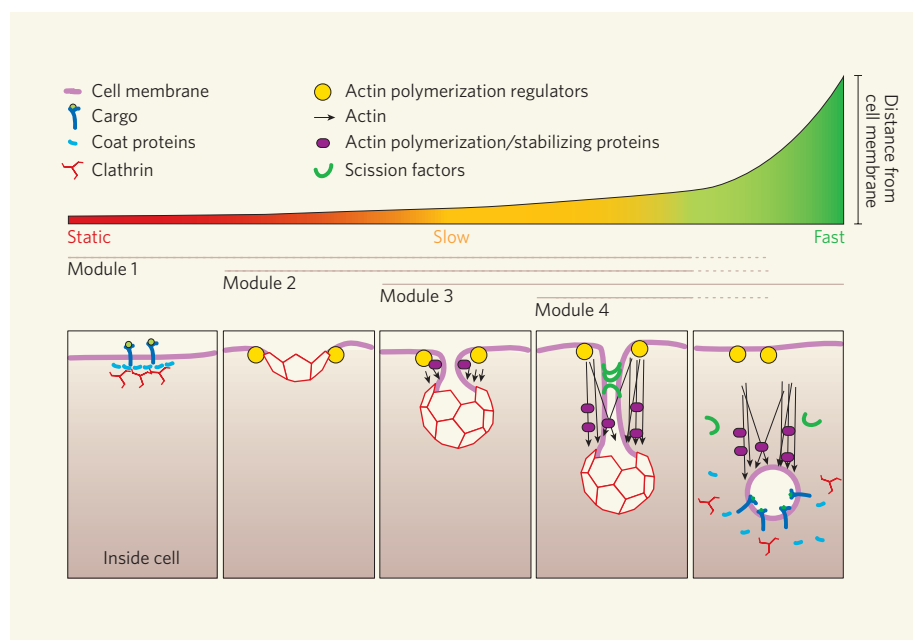


Figure 1 | The dynamics of endocytic vesicle formation. Initially, at the outer membrane of the cell, an immobile coat module assembles from clathrin and other cytoplasmic proteins (module 1). The coat captures cargo through interactions with the intracellular regions of the cargo proteins. Next, actin polymerization regulators (module 2) and actin polymerization/stabilizing modules associate with the coat, leading to the formation of actin filaments that drive slow inward movement of the coat (module 3). Finally, a scission module is recruited to separate the budded vesicle from the membrane (module 4). The freed vesicle moves rapidly into the cell and the coat components dissociate from the vesicle for additional rounds of vesicle formation.

(Fig. 1). Endocytosis begins with assembly at the plasma membrane of a clathrin-coat module containing proteins involved in cargo recruitment and coat formation. In cells lacking clathrin, the other coat components still assemble and their kinetic behaviours are only modestly affected. But in these cells the number of endocytic sites on the plasma membrane is severely reduced, implying that clathrin facilitates the assembly of the endocytic machinery.

The second module associates with the coat module and consists of a group of proteins involved in regulation of actin polymerization. Actin is a small protein that polymerizes into filaments; filament polymerization can generate the force necessary to move proteins and membranes.

Proteins that promote actin filament formation, accelerate filament assembly and stabilize new filaments make up the third module. Once this module assembles, stable actin filaments become apparent and the endocytic patch on the membrane begins to move, probably indicating membrane invagination (but this remains to be demonstrated). Treatment with a drug that abolishes actin polymerization, or genetic deletion of third-module components, trapped the endocytic patches at the plasma membrane, so it seems that actin polymerization is involved in vesicle invagination.

The fourth module appears transiently and contains proteins that, when purified, can constrict spherical membrane structures into

tubules. Remarkably, in cells lacking these proteins, the endocytic patch began to jut into the cell but sometimes snapped back to the cell surface. The authors suggest that this behaviour indicates a defect in vesicle release, consistent with membrane-constricting activity of the proteins. Once the vesicle is released, the coat protein module disassembles and the actin network continues to drive the vesicle deeper into the cell.

How well do the events in yeast correspond to those in mammalian cells? Although many components of the endocytic machinery are evolutionarily conserved between the two, there are differences in their requirements for actin and clathrin. In yeast, endocytosis is strictly dependent on actin assembly, but overall can proceed slowly without clathrin⁶. In contrast, clathrin-mediated endocytosis in mammalian cells seems to rely more on clathrin than on actin⁴. The results from yeast indicate that clathrin is an integral component of endocytic coats^{2,7}, with a significant but non-obligatory role in coat assembly. This confirms a long-standing observation that inactivation of clathrin causes immediate but partial endocytic defects⁸. Recent live-cell imaging of mammalian fibroblast cells revealed that, as in yeast, actin and actin-polymerizing proteins associate with most, if not all, clathrin coats, and promote invagination and vesicle movement^{9,10}. However, inhibition of actin polymerization causes only partial defects in endocytosis, indicating that although actin is important it is dispensable for endocytosis in

fibroblasts. So, overall, the fundamental features and components of clathrin-mediated endocytosis have been well conserved. The variation between yeast and mammalian cells probably reflects cell-type-specific requirements. For example, the higher internal pressure in yeast might present an energy barrier to vesicle formation that makes actin force-generating mechanisms more significant than in some mammalian cells.

The ability to visualize single vesicles forming in living cells and to perturb the process by gene inactivation provides an unprecedented opportunity to probe the mechanism of endocytosis at a molecular level. Drubin and colleagues have made a good start, but there are still many questions to be answered. How are endocytic sites selected, for example, or are they randomly initiated? What regulatory mechanisms ensure ordered progression from coat assembly to vesicle release and coat disassembly? How is actin polymerization triggered and then harnessed to drive membrane invagination and perhaps scission? And how is endocytosis coordinated with other cellular processes? It is evident that the dance has just begun.

Mara C. Duncan and Gregory S. Payne are in the Department of Biological Chemistry, David Geffen School of Medicine at UCLA, Los Angeles, California 90095, USA.
e-mail: mduncan@mednet.ucla.edu

1. Conner, S. D. & Schmid, S. L. *Nature* **422**, 37–44 (2003).
2. Kaksonen, M., Toret, C. P. & Drubin, D. G. *Cell* **123**, 305–320 (2005).
3. Roth, T. F. & Porter, K. R. *J. Cell Biol.* **20**, 313–332 (1964).
4. Perrais, D. & Merrifield, C. J. *Dev. Cell* **9**, 581–592 (2005).
5. Kaksonen, M., Sun, Y. & Drubin, D. G. *Cell* **115**, 475–487 (2003).
6. Baggett, J. J. & Wendland, B. *Traffic* **2**, 297–302 (2001).
7. Newpher, T. M., Smith, R. P., Lemmon, V. & Lemmon, S. K. *Dev. Cell* **9**, 87–98 (2005).
8. Tan, P. K., Davis, N. G., Sprague, G. F. & Payne, G. S. *J. Cell Biol.* **123**, 1707–1716 (1993).
9. Merrifield, C. J., Perrais, D. & Zenisek, D. *Cell* **121**, 593–606 (2005).
10. Yarus, D., Waterman-Storer, C. M. & Schmid, S. L. *Mol. Biol. Cell* **16**, 964–975 (2005).

Correction

In "Chemical biology: Bring them back alive" by Michael Yarus (*Nature* **438**, 40; 2005), the references were jumbled. The reference list should read as follows:

1. Fusz, S., Eisenführ, A., Srivatsan, S. G., Heckel, A. & Famulok, M. *Chem. Biol.* **12**, 941–950 (2005).
2. Yarus, M. *Annu. Rev. Biochem.* **74**, 179–198 (2005).
3. Wilson, D. S. & Szostak, J. W. *Annu. Rev. Biochem.* **68**, 611–647 (1999).
4. Lilley, D. M. *J. Curr. Opin. Struct. Biol.* **15**, 313–323 (2005).
5. Tarasow, T. M., Tarasow, S. L. & Eaton, B. E. *Nature* **389**, 54–57 (1997).
6. Helm, M., Petermeier, M., Ge, B., Fiammengio, R. & Jäschke, A. *J. Am. Chem. Soc.* **127**, 10492–10493 (2005).
7. Yarus, M. & Knight, R. in *The Genetic Code and the Origin of Life* (ed. de Pouplana, L.) 75–91 (Landes Bioscience, Georgetown, Texas, 2004).
8. Lawrence, M. S. & Bartel, D. P. *RNA* **11**, 1173–1180 (2005).
9. Yarus, M. *Cold Spring Harb. Symp. Quant. Biol.* **66**, 207–215 (2001).

BRIEF COMMUNICATIONS

Fruit bats as reservoirs of Ebola virus

Bat species eaten by people in central Africa show evidence of symptomless Ebola infection.

The first recorded human outbreak of Ebola virus was in 1976, but the wild reservoir of this virus is still unknown¹. Here we test for Ebola in more than a thousand small vertebrates that were collected during Ebola outbreaks in humans and great apes between 2001 and 2003 in Gabon and the Republic of the Congo. We find evidence of asymptomatic infection by Ebola virus in three species of fruit bat, indicating that these animals may be acting as a reservoir for this deadly virus.

Human Ebola outbreaks that occurred between 2001 and 2005 in Gabon and the Republic of the Congo were linked to concurrent outbreaks that devastated local gorilla and chimpanzee populations^{2,3}. To identify the viral reservoir, we undertook three trapping expeditions in areas close to infected gorilla and chimpanzee carcasses, just after their discovery (Fig. 1a). In total, 1,030 animals were captured, including 679 bats, 222 birds and 129 small terrestrial vertebrates, and were tested for evidence of infection by Ebola virus (for details, see supplementary information).

Of the infected animals identified during these field collections, immunoglobulin G (IgG) specific for Ebola virus was detected in serum from three different bat species (4 of 17 *Hypsignathus monstrosus*, 8 of 117 *Epomops franqueti* and 4 of 58 *Myonycteris torquata*). Two of the principal organs targeted by Ebola virus are the liver and spleen⁴. Viral nucleotide sequences were detected in these organs in other bats from the same populations (4 of 21, 5 of 117 and 4 of 141, respectively). No viral RNA was detected in kidney, heart or lung in these animals after amplification by polymerase chain reaction (PCR) and no viral nucleotide sequences were revealed in any of the other animal species tested.

Nucleotide-sequence analysis of purified PCR products identified seven different fragments amplified from the 13 PCR-positive animals, all clustering phylogenetically within the Zaire clade (Fig. 1b). The fragments differed not only from one collection to another, but also within a given collection, among the three bat species, and within a given species. The need to use nested PCR indicated that the viral RNA load in tissues was extremely low, which probably explains why we failed to isolate the virus itself.

Surprisingly, none of the IgG-positive animals was PCR-positive, and none of the PCR-positive animals was IgG-positive. This may

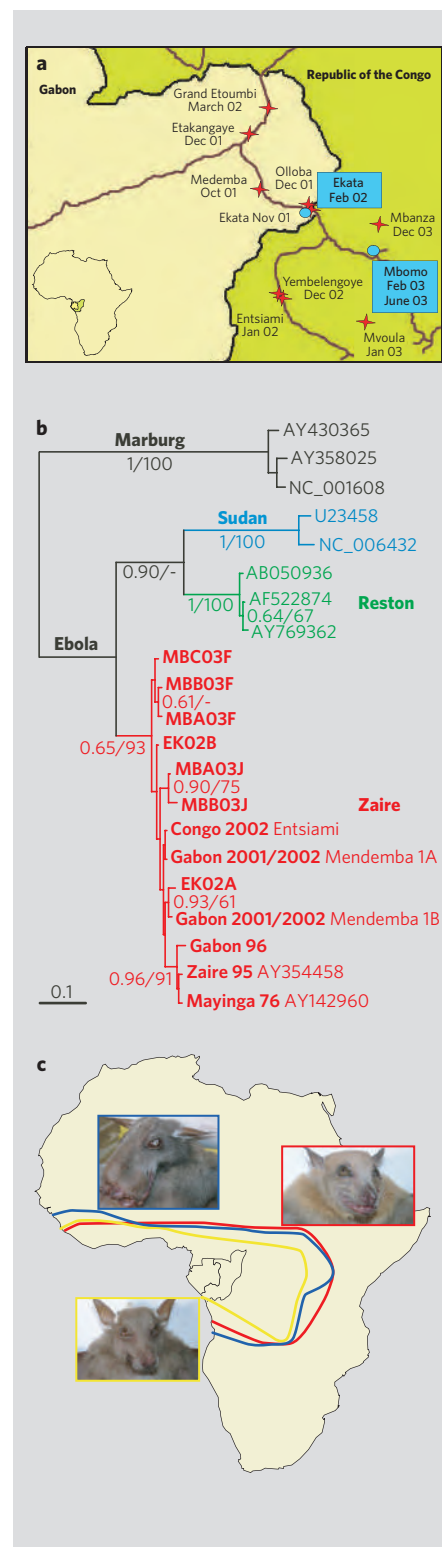
be because PCR-positive bats were recently infected and were tested before they developed a detectable immune response. Alternatively, it could be that differences in the virulence of Ebola virus strains led to different immunological responsiveness and viral replication patterns. Of the bat species collected at Mbomo in February 2003, 7 of 31 (22.6%) and 0 of 10 (0%) were PCR-positive and IgG-positive, respectively, but five months later the corresponding results were 4 of 184 (2.2%) and 12 of 160 (7.5%). These opposite trends in the PCR and serological results are consistent with the first hypothesis.

Each of the three bat species has a broad geographical range that includes regions of Africa where human Ebola outbreaks occur⁵ (Fig. 1c). Our findings support results of previous investigations that identify bats as candidate reservoirs for Ebola and Marburg viruses^{1,6}, and as reservoirs for the virus families *Paramyxoviridae* and *Rhabdoviridae*^{7–9}, which are genetically related to Ebola.

Mortality among great apes from Ebola infection can increase during the dry seasons³ when fruit is scarce in the forest — conditions that foster contact between animals as they compete for food. Immune function in bats also changes during these periods¹⁰, for example as a result of food scarcity or pregnancy, which would favour viral replication and — aided by aggressive interactions — increase infection among great apes. These factors may contribute to the episodic nature of Ebola outbreaks.

Although other bat and animal species may also act as Ebola virus reservoirs, insight into

Figure 1 | Fruit bats as potential carriers of Ebola virus. **a**, Dates and locations of animal-trapping sites (blue) and of Ebola virus outbreaks among humans (red stars) in Gabon and the Republic of the Congo. **b**, Phylogeny of Ebola viruses inferred from RNA polymerase sequences. Values below branches are bayesian posterior probabilities (left of slash; values less than 0.5 not shown); bootstrap percentages were obtained by maximum parsimony (right of the slash; values under 50% not shown). (GenBank accession numbers, DQ 205409–205415.) Sequences of the subtype Zaire (red) share five nucleotide signatures in positions 1,755 (T), 1,800 (G), 1,857 (T), 2,002 (A) and 2,003 (C) of the complete coding sequence of the gene encoding RNA polymerase. **c**, Geographic distribution (inside coloured lines) of the fruit bats *Hypsignathus monstrosus* (blue), *Epomops franqueti* (red) and *Myonycteris torquata* (yellow).



the behavioural ecology of the bat species identified here should help to improve protection of the great apes from Ebola virus. Human infection directly from fruit bats might in part be countered by education, as these animals are eaten by local populations living in the outbreak regions.

Eric M. Leroy*, **Brice Kumulungui***,
Xavier Pourrut*, **Pierre Rouquet***,
Alexandre Hassanin†, **Philippe Yaba***,
André Délicat*, **Janusz T. Paweska‡**,
Jean-Paul Gonzalez§, **Robert Swanepoel‡**

*Centre International de Recherches Médicales de Franceville, BP 769 Franceville, Gabon
e-mail: eric.leroy@ird.fr

†Muséum National d'Histoire Naturelle, UMR 5202, Paris 75005, France

‡National Institute for Communicable Diseases, Special Pathogens Unit, Private Bag X4,

Sandringham 2131, South Africa
§Institut de Recherche pour le Développement, UR178, Mahidol University at Salaya, Nakhonpathom 73170, Thailand

1. Pourrut, X. *et al. Microbes Infect.* **7**, 1005–1014 (2005).
2. Walsh, P. D. *et al. Nature* **422**, 611–614 (2003).
3. Leroy, E. M. *et al. Science* **303**, 387–390 (2004).
4. Geisbert, T. W. *et al. Am. J. Pathol.* **163**, 2347–2370 (2003).
5. Bergmans, W. *Beaufortia* **39**, 89–152 (1989).
6. Swanepoel, R., Leman, P. A. & Burt, F. J. *Emerg. Infect. Dis.* **2**, 321–325 (1996).
7. Sanchez, A., Killey, M. P., Klenk, H.-D. & Feldmann, H. J. *Gen. Virol.* **73**, 347–357 (1992).
8. Halpin, K., Young, P. L., Field, H. E. & Mackenzie, J. S. J. *Gen. Virol.* **81**, 1927–1932 (2000).
9. Arguin, P. M. *et al. Emerg. Infect. Dis.* **8**, 258–262 (2002).
10. Langevin, P. & Barklay, R. M. R. *Mammal. Species* **357**, 1–4 (1990).

Supplementary information accompanies this communication on Nature's website.

Competing financial interests: declared none.
doi:10.1038/438575a

PALAEOECOLOGY

A gigantic fossil arthropod trackway

A unique, complex trackway has been discovered in Scotland: it was made roughly 330 million years ago by a huge, six-legged water scorpion that was about 1.6 m long and a metre wide. To my knowledge, this is not only the largest terrestrial trackway of a walking arthropod to be found so far, but is also the first record of locomotion on land for a species of *Hibbertopterus* (Eurypterida). This evidence of lumbering movement indicates that these giant arthropods, now extinct, could survive out of water at a time when the earliest tetrapods were making their transition to the land.

The trackway (Fig. 1a, b) is exposed on a bedding plane close to the base of a sandstone section in a non-marine sequence. It is 6 m long, 0.90–0.98 m wide and consists of sinuous, paired belts of appendage prints flanking a sub-central groove. The trace-maker had at least three pairs of appendages of different lengths (heteropodous), which moved in phase. The longest, outer limbs left elongated crescent-shaped prints (series A in Fig. 1b, green), which overlap slightly or coalesce into a linked series of arcs. The stride length is therefore less than the series-A print length (average, 0.27 m) and indicates that the animal was crawling extremely slowly.

Lines of elongate, crescentic or sigmoidal prints (series B in Fig. 1b, blue) lie inside series A, and further elliptical prints (series C in Fig. 1b, yellow), made by the shortest appendages, can be detected inside these. In places, the series-C prints have been erased by the central groove, which was made by the posterior part of the animal. This is trapezoidal in cross-section and its base is deeper at the margins and slightly raised in the centre. Occasional oblique lineations on the sides and base of the groove indicate that the motion

was jerky. The sinuous curve of the groove is smaller in amplitude than, and out of phase with (by about 0.5–0.6 m), the trackway margins, which reveals the direction of locomotion (Fig. 1a, b). The slow, stilted progression, together with the dragging of the posterior, indicates that the animal was not buoyant and that it was probably moving out of water.

There are several groups of Lower Carboniferous (Asbian) arthropods that might have been capable of leaving large trackways¹, but only the water scorpions, or eurypterids¹, are likely to have left the trackway described here. The pattern and character of the limb prints is most consistent with a relatively short-limbed and markedly heteropodous hibbertopteroid eurypterid^{2–5} (Fig. 1c). The double-keeled underside of the terminal tail plate of these animals^{4,5} matches the character of the central groove.

Fragmentary exoskeletal remains of *Hibbertopterus* and related forms are relatively well known from Scottish Lower Carboniferous rocks^{2–5} and were first described from West Lothian in 1831 (ref. 2). The trackway-maker (Fig. 1c) would have been comparable in size to the largest known hibbertopteroid body fossils, which have head shields^{3–5} that are 0.65 m wide.

The short length of the relative stride in the trackway emphasizes the extreme slowness of the gait and differentiates it from other eurypterid trackways within the ichnogenus *Palmichnium*^{1,6–9}. This trace is 0.2 m (25%) wider than any other trackway of this type¹. The only larger known invertebrate trackway, although also attributed to a eurypterid, is very different in character and appears to have been made by a swimming animal¹⁰.

Martin A. Whyte

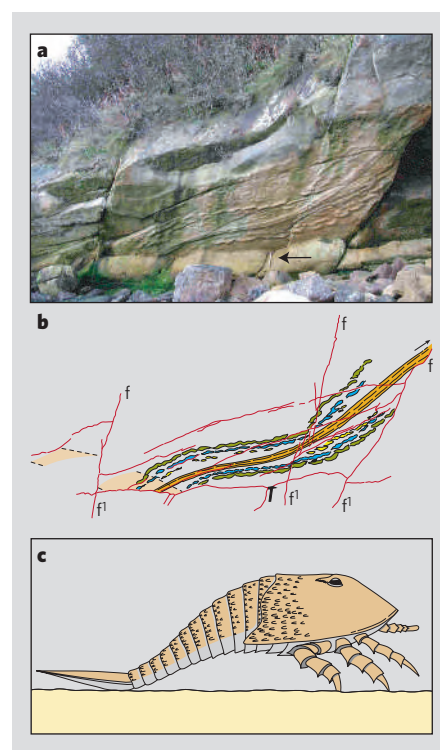


Figure 1 | Hibbertopteroid trackway from Lower Carboniferous (Asbian) rocks in Scotland.

a, View of the trackway on the undersurface of an overhanging sandstone bed, which is dipping at 45° away from the viewer. The hammer (arrowed) in the photograph is 30 cm long, but the oblique view affects scale and relative proportions.

b, Interpretive diagram showing track features, position of a second, smaller (0.80 m wide) trackway and the position in the rock of microfaults (f-f'), joints and bedding traces (red lines). Arrow indicates movement direction of the animal. Trackway: orange, central groove; series A, B and C are shown in green, blue and yellow, respectively. **c**, Reconstruction of the hibbertopteroid eurypterid trackway-maker. This arthropod was about 1.6 m. long (for clarity, the limbs on the left of the body are omitted).

Department of Geography, University of Sheffield, Brookhill, Sheffield S10 2TN, UK
e-mail: m.a.whyte@sheffield.ac.uk.

1. Briggs, D. E. G. & Rolfe, W. D. I. *J. Paleontol.* **57**, 377–390 (1983).
2. Waterston, C. D. *Trans. R. Soc. Edinb.* **61**, 265–288 (1957).
3. Waterston, C. D. *Trans. R. Soc. Edinb.* **68**, 1–20 (1968).
4. Jeram, A. J. & Selden, P. A. *Trans. R. Soc. Edinb.* **E 84**, 301–308 (1994).
5. Selden, P. A., Corronca, J. A. & Hünicken, M. A. *Biol. Lett.* **1**, 44–48 (2005).
6. Braddy, S. J. & Milner, A. E. *Can. J. Earth Sci.* **35**, 1116–1122 (1998).
7. Braddy, S. J. & Almond, J. E. *J. Afr. Earth Sci.* **29**, 165–177 (1999).
8. Draganits, E., Braddy, S. J. & Briggs, D. E. G. *Palaeos* **16**, 126–147 (2001).
9. Hanken, N. M. & Stormer, L. *Foss. Strata* **4**, 255–270 (1975).
10. Almond, J. E. *Geobulletin* **45**, 28 (2002).

Competing financial interests: declared none.
doi:10.1038/438576a

BRIEF COMMUNICATIONS ARISING online
♦ www.nature.com/bca see Nature contents.

GEOCHRONOLOGY

Age of Mexican ash with alleged 'footprints'

Arising from: R. Dalton *Nature* doi: 10.1038/news050704-4 (2005)

A report of human footprints preserved in 40,000-year-old volcanic ash near Puebla, Mexico (<http://www.royalsoc.ac.uk/exhibit.asp?id=3616&tip=1>), was the subject of a press conference that stirred international media attention¹. If the claims (www.mexicanfootprints.co.uk) of Gonzalez *et al.* are valid, prevailing theories about the timing of human migration into the Americas would need significant revision. Here we show by $^{40}\text{Ar}/^{39}\text{Ar}$ dating and corroborating palaeomagnetic data that the basaltic tuff on which the purported footprints are found is 1.30 ± 0.03 million years old. We conclude that either hominid migration into the Americas occurred very much earlier than previously believed, or that the features in question were not made by humans on recently erupted ash.

We report the results of $^{40}\text{Ar}/^{39}\text{Ar}$ dating and palaeomagnetic analysis of volcanic rock (tephra) from the surface carrying the indentations interpreted by Gonzalez *et al.* as footprints. The tephra is a moderately indurated olivine basalt lapilli tuff, in which the lapilli are cemented by a matrix composed of fine-grained claylike minerals that seem to be formed by marginal alteration of the lapilli. At the sampled locality in Toloquilla quarry ($18^\circ 55.402' \text{ N}$, $98^\circ 009.375' \text{ W}$), this tuff occurs near the top of a stratigraphic succession of at least several metres of similar, rhythmically bedded tuff layers, 1–10 cm thick, which are exposed in small quarrying trenches and pits.

Nine samples (six comprising a single 2–4-mm lapillus, three comprising 3–11 such

lapilli) were analysed by incremental laser-heating $^{40}\text{Ar}/^{39}\text{Ar}$ analysis. Isotopic data are available at www.bgc.org/renne/data/XalneneArData.xls. All the aliquots yielded mutually indistinguishable plateau ages ranging from 1.26 ± 0.10 to 1.47 ± 0.30 Myr (Fig. 1). A weighted mean of the nine plateau ages is 1.30 ± 0.03 (2 σ) Myr.

The palaeomagnetic analysis (Fig. 2) reveals two components of essentially opposite polarity, a reverse component held by magnetite within the lapilli and an antipodal normal component held by goethite within the claylike matrix. The sample is azimuthally unoriented but the stratigraphic top and bottom are known. The reverse polarity component held by the tuff is thermoremanent, which strongly suggests that the tuff pre-dates the Brunhes/Matuyama geomagnetic polarity transition at about 790 kyr (ref. 2). A reverse polarity is consistent with magnetization acquired during chron C1r.2r, between 1.07 and 1.77 Myr (ref. 3), as indicated by the $^{40}\text{Ar}/^{39}\text{Ar}$ data. Because only the *in situ* top of the sample

is known, the declination is arbitrary but the inclination (-32.1°) is meaningful and compares favourably with that (-33.4°) expected from a reversed geocentric axial dipole at this latitude.

The clustering of directions from the multi-lapilli specimens indicates that the lapilli were emplaced hot, above the blocking temperature of their constituent titanomagnetite grains, and have not been disturbed since emplacement. If the lapilli in the sample were derived by erosion of pre-existing older tephra, their age could substantially pre-date deposition. This possibility can be ruled

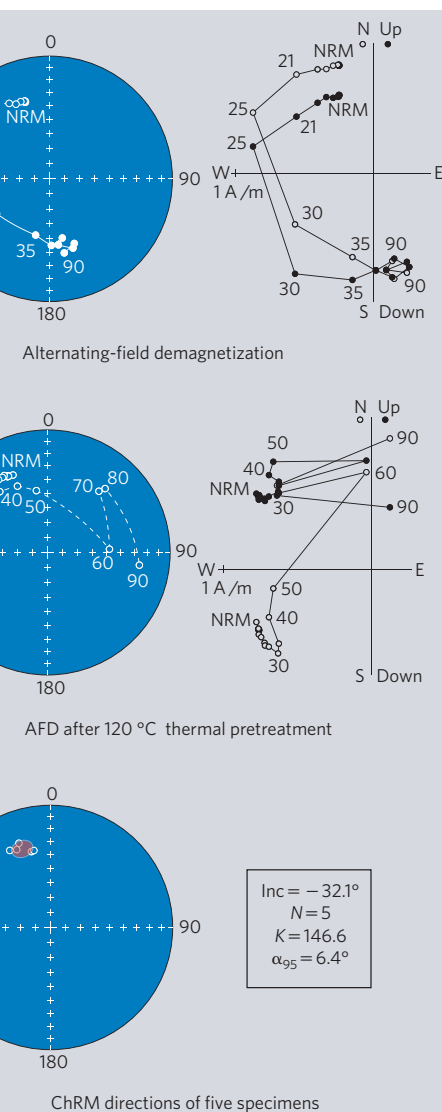


Figure 2 | Equal area and orthogonal demagnetization diagrams for a specimen of volume 16 cm³. **a**, Only alternating-field demagnetization (AFD) used; and **b**, AFD used after a 30-min demagnetization step at 120 °C. The heating step removes a secondary component due to goethite (Néel temperature, 120 °C; ref. 7). NRM, natural remanent magnetization. **c**, Reversed components from five such samples were determined using least-squares methods⁸ and combined as a fisherian characteristic direction (ChRM) with the confidence ellipse shown. K, concentration parameter; α_{95} , 95% confidence interval.

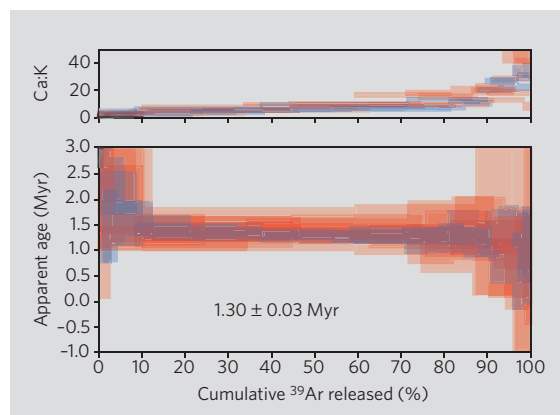


Figure 1 | Analysis of Xalnene ash samples from central Mexico. Top, calcium:potassium spectra show effects of degassing calcic minerals (clinopyroxene and plagioclase cores) at increasing temperature. Bottom, age spectra for nine samples, degassed in 5–13 steps with a carbon dioxide laser, are superimposed and shown with the weighted mean of nine plateau ages determined in each of the individual experiments. Ages are reported relative to the 1.194-Myr Alder Creek sanidine standard⁶. Single lapilli samples are shown in red, multi-lapilli samples in blue.

out by the homogeneous textural, compositional and geochronological characteristics of the lapilli, paucity of extraneous volcanic clasts, absence of mechanical abrasion features, lack of soil development or sedimentary horizons within the Xalnene tuff sequence, and palaeomagnetic evidence for hot emplacement.

If the markings on the exposed surface of the tuff are human footprints recorded soon

after its eruption, the obvious implication is that they are 1.3 million years old. This would be truly extraordinary as such antiquity pre-dates even the most speculative credible inferences about the first known appearance of *Homo sapiens* in the western hemisphere by more than a million years. Indeed, the Xalnene tuff pre-dates the first known appearance of *H. sapiens* (in Africa^{4,5}) by more than a million years. If the markings are hominid footprints, they would be most likely to have been made by a hominid that existed before *H. sapiens*, and we consider such a possibility to be extremely remote. We conclude that the identification of any of these features as syn-depositional hominid footprints is erroneous.

Paul R. Renne*†, **Joshua M. Feinberg**†,

Michael R. Waters‡, **Joaquin Arroyo-Cabres**§, **Patricia Ochoa-Castillo**||, **Mario Perez-Campa**¶, **Kim B. Knight**†

*Berkeley Geochronology Center, Berkeley, California 94709, USA
e-mail:prene@bgc.org

†Department of Earth and Planetary Science, University of California, Berkeley, California 94720, USA

‡Departments of Anthropology and Geography, and Center for the Study of the First Americans, Texas A&M University, College Station, Texas 77843-4352, USA

§Laboratório de Arqueozoologia, Subdirección de Laboratórios y Apoyo Académico, Instituto Nacional de Antropología e Historia, CP 06060, México

||Subdirección de Arqueología, Museo Nacional de Antropología, Reforma y Gandhi s/n, CP 11560, México

¶Proyecto Cuicuilco, Instituto Nacional de Antropología e Historia, CP 06700, México

1. Dalton, R. *Nature* doi:10.1038/news050704-4 (2005).
2. Coe, R. S., Singer, B. S., Pringle, M. S. & Zhao, X. *Earth Planet. Sci. Lett.* **222**, 667–684 (2004).
3. Cande, S. C. & Kent, D. V. *J. Geophys. Res.* **100**, 6093–6095 (1995).
4. Clark, J. D. *et al. Nature* **423**, 747–752 (2003).
5. McDougall, I., Brown, F. H. & Fleagle, J. G. *Nature* **433**, 733–736 (2005).
6. Nomade, S. *et al. Chem. Geol.* **318**, 315–338 (2005).
7. Dunlop, D. J. & Özdemir, Ö. *Rock Magnetism: Fundamentals and Frontiers* (Cambridge University Press, 1997).
8. Kirschvink, J. L. *Geophys. J. R. Astron. Soc.* **62**, 699–718 (1980).

doi:10.1038/nature04425

**Cover illustration**

Micrograph of a central nervous system synapse, courtesy of John Heuser (Washington University) and Thomas Reese (National Institutes of Health).

Editor, Nature

Philip Campbell

Insights Publisher

Sarah Greaves

Insights Editor

Lesley Anson

Production Editor

Maria Hodges

Senior Art Editor

Martin Harrison

Art Editor

Nik Spencer

Layouts

Nadia Rooney

Sponsorship

Claire Hines
Claudia Banks

Production

Sue Gray

Marketing

Robin Brown

Editorial Assistant

Laura Shaw

The Macmillan Building,

4 Crinan St

London N1 9XW, UK

Tel: +44 (0) 20 7833 4000

e-mail: nature@nature.com



nature publishing group

MEMBRANE BIOLOGY

Biological membranes surround and compartmentalize cells. They form the interface between the cell and its environment, and are key players in cellular homeostasis and metabolic-energy transduction. This Insight provides a flavour of the many facets of contemporary membrane biology.

Biological membranes are more complex than was first thought when the 'fluid mosaic model' was proposed in 1972. Membranes are a mixture of many different types of lipidic and protein components, and their relative amounts and composition differ between functionally distinct domains. There is, however, some fluidity between these membrane compartments as vesicles bud off from one compartment and fuse with another. Ongoing studies are revealing the protein and lipidic signals that control the exquisite specificity of these dynamic vesicular processes.

Membrane proteins typically make up around a third of the proteome of a cell. Nevertheless, molecular-level understanding of membrane proteins lags far behind that of water-soluble proteins owing to the difficulty in obtaining high-resolution structural information. The recent and rapidly expanding crop of membrane protein structures is revolutionizing understanding of the principles that govern the folding of these proteins.

With a range of topics from membrane protein biophysics to the cell biology of membrane processes, this Insight introduces some of the most exciting current research in this field. We are indebted to all the authors who contributed.

We are pleased to acknowledge the financial support of Avanti Polar Lipids, which contributed towards the distribution of this Insight. As always, *Nature* retains sole responsibility for editorial content and peer review.

Deepa Nath, Senior Editor

INTRODUCTION

578 Membranes are more mosaic than fluid

D. M. Engelman

REVIEWS

581 Solving the membrane protein folding problem

J. U. Bowie

590 Membrane curvature and mechanisms of dynamic cell membrane remodelling

H. T. McMahon &
J. L. Gallop

597 Organelle identity and the signposts for membrane traffic

R. Behnia & S. Munro

605 Plasma membrane phosphoinositide organization by protein electrostatics

S. McLaughlin &
D. Murray

612 Role of cholesterol and lipid organization in disease

F. R. Maxfield & I. Tabas

nature
insight

Membranes are more mosaic than fluid

Donald M. Engelman¹

The wealth of new data on membrane protein structures and functions is changing our general view of membrane architecture. Some of the key themes that are emerging are that membranes are patchy, with segregated regions of structure and function, that lipid regions vary in thickness and composition, and that crowding and ectodomains limit exposure of lipid to the adjacent aqueous regions.

Given their biological importance, membranes have been surprisingly neglected by biochemists until recently. Perhaps this is understandable in view of the technical hurdles that working with them presents. Most methods require purification and observation in aqueous environments alien to the molecular design of a membrane, and so the field had to rely on oversimplified views that still dominate the texts and teaching in this area. But now we have a rising number of high-resolution structures, an abundance of functional data and an evolving conceptual basis for framing more pointed questions. This is leading to a great expansion of interest in the area. Articles in this Insight expose current views of the importance, findings and concepts in membrane biology in some regions of the emerging landscape.

The reductionist view of biology, to which many adhere, rests in part on the structure–function hypothesis: that the structures we find are there for specific functional reasons selected by evolution. In the case of membranes, we might start with the origin of life, noting that compartmentalization is essential for an organism, and that with compartmentalization must come specific ways to surmount the barrier defining the boundary of the compartment — the membrane. Thus, the lipid bilayer, which spontaneously forms permeability barriers surrounding aqueous interiors, must be modified by macromolecules for the uptake of nutrients and the disposal of waste. Further refinements led to the use of the barrier for its energy-storage properties and to the creation of ways to pass information between a cell and its environment.

To frame a context for the reviews that follow, a few general perspectives are presented briefly below. To develop these themes fully would require a much longer text (perhaps a book?), so only representative references are given here, and use is made of the references in the longer treatments by the other authors.

Patchiness in the membrane plane

An influential step in the study of membranes was taken with the development by Singer and Nicholson in 1972 of the ‘fluid mosaic model’¹, which pulled together findings and ideas from the preceding decade. The model has become the standard conceptualization of membrane architecture and is shown redrawn in Fig. 1a as it appears in virtually all biochemistry texts. As important and insightful as this model has been, the emergence of new findings during the passage of 33 years has weakened the generalizations it contains, and it is now appropriate to examine some of them more closely. The model includes the ideas that the proteins of a membrane are dispersed, are at low concentration and that they match the hydrophobic dimension of an unperturbed lipid bilayer with peripheral belts of exposed hydrophobic side chains. The lipid is seen as a sea in which mainly monomeric proteins float unencumbered, and the bilayer surface is exposed directly to the aqueous environment.

Each of these ideas is misleading. Most of the authors of the following reviews write of the preferential associations of molecules in the membrane plane, and as an introduction I suggest that such associations are expected, that membranes are typically crowded and that their bilayers vary considerably in thickness.

Is a membrane a random two-dimensional liquid? In the Singer–Nicholson model, molecules are distributed randomly in two dimensions. But we know from first principles and from experimental observation that non-randomness is the rule. Consider a mixture of *n*

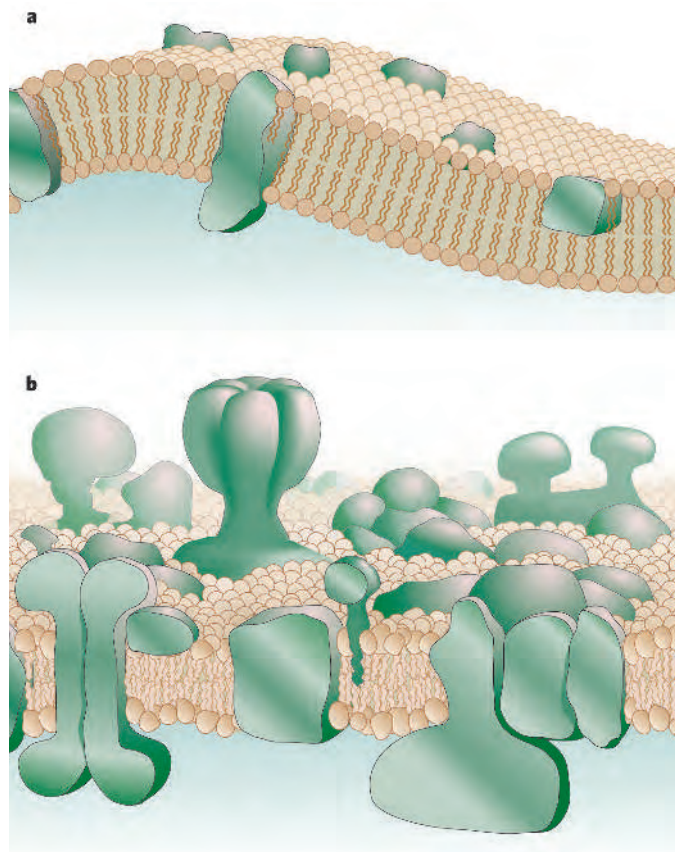


Figure 1 | General models for membrane structure. **a**, The Singer–Nicholson ‘fluid mosaic model’ (ref. 1). **b**, An amended and updated version.

¹Department of Molecular Biophysics and Biochemistry, Yale University, Box 208114, New Haven, Connecticut 06520-8114, USA.

lipid and protein components in a membrane. The planar distribution can be random only if all pairwise interaction energies of the n different molecular species are within thermal energies (about $0.6 \text{ kcal mol}^{-1}$ at room temperature) of each other. In a plasma membrane there are many species of lipids and proteins. The *Escherichia coli* genome, for example, codes for more than a thousand putative helical transmembrane proteins², giving more than half a million kinds of pairwise combinations. A narrow range of interaction energies is a highly improbable condition given the range of known intermolecular interactions from hydrogen bonds, packing, electrostatics and the hydrophobic effect. Indeed, simply rotating a pair of identical helices against each other or changing a single interfacial side chain can result in interaction variations of several times kT (refs 3, 4). Thus, it should have been expected that regions of biased composition would exist and that the environments of proteins should vary, because it is highly improbable that interaction energies will match each other across all protein and lipid species in a membrane. Time-invariant complexes, transient associations and biased distributions should be the norm. Evolution, ever seeking to exploit the natural tendencies of molecules, has seized the opportunity to craft functional associations, and it is clear that there are functional protein complexes, separated lipid compositional areas and regions of functional specialization, although we do not yet know their extent.

Many experimental observations now support the patchiness of membranes as a principle. A clear case is that proteins are mainly oligomeric. In the thoughtful database compiled by White (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html), almost all of the unique structures are oligomeric, and many are hetero-oligomers. Such oligomers are formed using strong associations and are resistant to dissociation by detergent. The excellent folding discussion offered here by Bowie (see page 581) implies that the side-to-side helix associations guiding folding also guide oligomer formation. More associations should be revealed when better means are found to extract them from or observe them in a membrane, and improved methods that reveal larger complexes are appearing⁵. Views of the participation of membrane proteins in organizing large functional complexes are beginning to emerge⁶. In common with proteins, lipids also tend to group together, forming lipid–lipid and lipid–protein interactions. Many lipids are seen in crystal structures to form specific complexes with proteins, most famously in the only structure of an entire membrane that we know — the purple membrane from *Halobacterium salinarum*⁷. A large body of literature shows that lipids on their own form regions of separated composition in the plane of pure lipid vesicles, as discussed here in the review by Maxfield and Tabas (p. 612), who also examine the role of lipids in disease. The ongoing discussion of ‘rafts’ is a case in point⁸. Further, distortions of the membrane thickness through lipid–protein interactions will create strained regions, as argued below. So, it would seem that patchiness is the order of the day.

Functional patchiness underlies ideas developed in three of the reviews that follow. McLaughlin and Murray (p. 605) discuss the idea of spatial organization imposed by electrostatic interactions as a way to understand the diverse functions of bisphosphatidylinositol (PIP₂). The discussion of the requirements created by the need for organelle identity, maintenance and function, presented by Behnia and Munro (p. 597), also implies membrane regionalization. It could further be argued that the functional correlates of the membrane curvature discussed by McMahon and Gallop (p. 590) depend on planar segregation of membrane contents and of curvature-inducing proteins when vesicles are formed.

Membrane thickness

What determines the thickness of a membrane lipid bilayer? The fluid mosaic model posits that “the structures of the lipid in the membrane and of the lipid in isolated aqueous dispersion are closely similar” and that “hydrophobic and hydrophilic interactions are to be maximized and the lowest free energy state is to be attained for the intact membrane in an aqueous environment”¹. It follows that membrane proteins

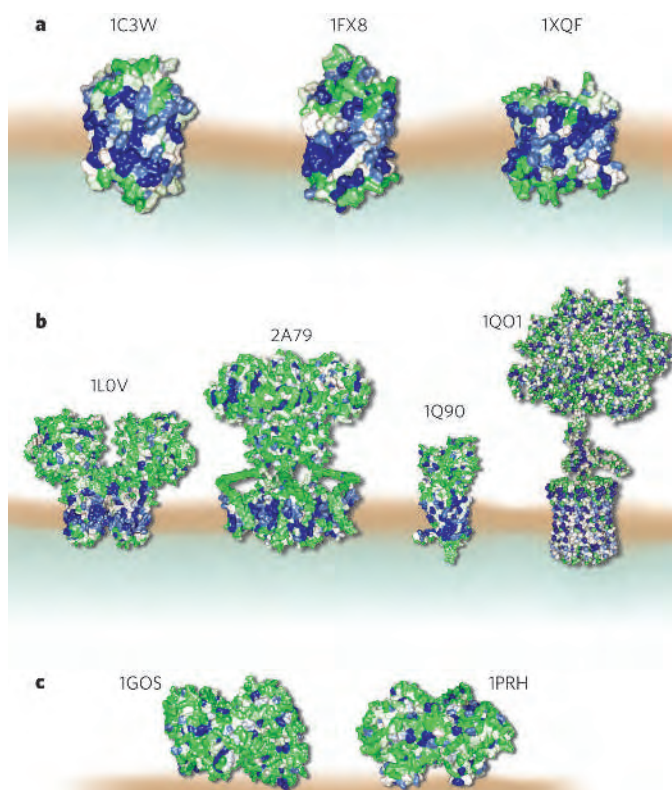


Figure 2 | Known structures for membrane proteins^{4, 14–28} represented using the energy scale of ref. 33. **a**, Proteins largely within the membrane bilayer. **b**, Proteins with large extramembrane regions. **c**, Proteins covering large regions of lipid³⁴. The green indicates amino acids with a favourable interaction with the hydrophobic lipid region, blue a favourable interaction with water. Calculations are by Jonathan Sachs and Michael Strickler.

should have evolved their hydrophobic regions to fit the lipid dimension perpendicular to the membrane plane, since the exposure of hydrophobic surface area to water is unfavourable by about $25 \text{ cal } \text{\AA}^{-2}$, and the exposure of, say, about $5 \text{ \AA}^2$ of bacteriorhodopsin surface out of the membrane would be unfavourable by about 25 kcal mol^{-1} . Inspection of the known structures shows that they vary in hydrophobic dimension around their peripheries and also from one to another. Something must give — either the protein distorts to match the bilayer or the lipid distorts to match the protein, or both. The fluidity of the lipid and the relative rigidity of the proteins⁹ suggest that it is the main lipid that distorts to match the protein, and both modelling and experiment support this view^{10,11}, although protein distortion may occur in extreme cases of protein–lipid mismatch¹².

If the lipid distorts to cover the hydrophobic area of a membrane protein, the transmembrane dimension of a bilayer in membranes with high protein-to-lipid ratios must be variable. Further, if the distortion is asymmetrical across a bilayer, curvature can result (discussed in the review by McMahon and Gallop, p. 590). Local distortion of the bilayer is likely to influence protein interactions; for example, the peripheral energy of distorting the bilayer may enhance interactions that reduce the peripheral contour length. What effect the energy of distorting the bilayer might have on the protein itself is not known, but might have functional relevance for cases where the protein varies its transmembrane conformation in the course of activity¹³.

Area occupancy by protein and lipid

How much of the membrane bilayer area is occupied by protein? In general, we do not know the answer to this question, yet the answer

strongly influences concepts of membrane organization. Further, proteins may occupy small areas at the bilayer level but have large ectodomains covering lipid and creating steric restrictions. Most drawings of the fluid mosaic model greatly exaggerate the lipid area in both senses — the area occupied by protein and the area covered by protein are shown as small and zero, respectively.

Membrane protein shapes vary greatly, as shown in Fig. 2. Some are largely contained within the bilayer^{4,14–18}, as in the examples in Fig. 2a. Many protein complexes, such as the ATP synthase, have large structures outside the lipidic region that will create steric contacts and other interactions outside the bilayer^{19–26} (Fig. 2b). These may occupy larger areas in projection onto the membrane than do the transmembrane regions. For example, the F1 ATPase ectodomains occupy about four times the membrane surface area of the transmembrane region²³. Proteins anchored by single helices or by lipidic anchors such as fatty acids can cover large regions of a membrane with protein surfaces^{27,28}, as in the examples shown in Fig. 2c. Interactions of the ectodomain structures are known to be functionally important in many cases, such as the tyrosine kinase receptors²⁹. Inspection of the examples in Fig. 2 shows that the exposure of membrane lipid surface may be rather small, for example when a plasma membrane is viewed from either the cytoplasm or the extracellular milieu. However, some proteins associate and dissociate with lipid as part of their function³⁰, so some lipid exposure must be maintained. Whether lipid exposure to the cytoplasm might be used to control or focus such interactions is at present unexplored.

Fluidity in the context of order

Although most membranes exhibit fluidity, with rapid diffusion observed for some lipid and protein species in the plane of a membrane, recent measurements using single-particle tracking reveal a complex set of restrictions on protein lateral mobility. These include directed motion, confined motion and anomalous diffusion³¹. Although the observation that some proteins can move relatively freely suggests that a subset may not be in larger assemblies, crowding, ectodomain collisions, transbilayer interactions, adhesion sites and cytoskeletal structure produce a variety of restrictions on the motion of most proteins and lipids³². Not yet considered are the additional constraints imposed by lipid–protein interactions through complex formation and thickness perturbation.

Fluidity must be reconciled with order. It follows that the patchiness of many membranes must be local enough for there to be channels of lipid separating regions of protein assemblies, but this constraint would still allow large segregated regions. The sizes and variability of segregated regions are still to be established.

Consequences for current concepts

The concepts developed above lead us to a view of a membrane that has variable patchiness, variable thickness and higher protein occupancy than has generally been considered. It will be a challenge to the immense and excellent body of work on pure lipid systems to assimilate the perturbations by proteins. Of course, there will be variability — myelin membranes are low in protein content; photosynthetic membranes high. But, while we await improved measurements, the modified view sketched in Fig. 1b is suggested as a guide to thinking, and as a context for the functional interactions that are discussed in the reviews that follow.

modulates the stability of a GxxxG-mediated transmembrane helix-helix dimer. *J. Mol. Biol.* **341**, 991–998 (2004).

4. Adams, P. D., Engelman, D. M. & Brunger, A. T. Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. *Proteins* **26**, 257–261 (1996).
5. Stenberg, F. et al. Protein complexes of the *Escherichia coli* cell envelope. *J. Biol. Chem.* **280**, 34409–34419 (2005).
6. Wong, W., Scott, J. D. AKAP signalling complexes: focal points in space and time. *Nature Rev. Mol. Cell Biol.* **5**, 959–970 (2004).
7. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* **259**, 393–421 (1996).
8. Brown, D. A. & London, E. Functions of lipid rafts in biological membranes. *Annu. Rev. Cell Dev. Biol.* **14**, 111–136 (1998).
9. Zaccai, G. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science* **288**, 1604–1607 (2000).
10. Petrache, H. I. et al. Hydrophobic matching mechanism investigated by molecular dynamics simulations. *Langmuir* **18**, 1340–1351 (2002).
11. Mitra, K., Ubarretxena-Belandia, I., Taguchi, T., Warren, G. & Engelman, D. M. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl Acad. Sci. USA* **101**, 4083–4088 (2004).
12. Williamson, I. M., Alvis, S. J., East, J. M. & Lee, A. G. Interactions of phospholipids with the potassium channel KcsA. *Biophys. J.* **83**, 2026–2038 (2002).
13. Perozo, E., Kloda, A., Cortes, D. M. & Martina, B. Physical principles underlying the transduction of bilayer deformation forces during mechanosensitive channel gating. *Nature Struct. Biol.* **9**, 636–637 (2002).
14. Abramson, J. et al. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301**, 610–615 (2003).
15. Fu, D. et al. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481–486 (2000).
16. Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* **291**, 899–911 (1999).
17. Zheng, L., Kostrewa, D., Berneche, S., Winkler, F. K. & Li, X. D. The mechanism of ammonia transport based on the crystal structure of AmtB of *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **101**, 17090–17095 (2004).
18. Abramson, J. et al. The structure of the ubiquinol oxidase from *Escherichia coli* and its ubiquinone binding site. *Nature Struct. Biol.* **7**, 910–917 (2000).
19. Iverson, T. M., Luna-Chavez, C., Cecchini, G. & Rees, D. C. Structure of the *Escherichia coli* fumarate reductase respiratory complex. *Science* **284**, 1961–1966 (1999).
20. Long, S. B., Campbell, E. B. & Mackinnon, R. Crystal structure of a mammalian voltage-dependent Shaker family K⁺ channel. *Science* **309**, 897–903 (2005).
21. Kurisu, G., Zhang, H., Smith, J. L. & Cramer, W. A. Structure of the cytochrome b₆f complex of oxygenic photosynthesis: tuning the cavity. *Science* **302**, 1009–1014 (2003).
22. Miyazawa, A., Fujiyoshi, Y. & Unwin, N. (2003). Structure and gating mechanism of the acetylcholine receptor pore. *Nature* **423**, 949–955.
23. Stock, D., Leslie, A. G. & Walker, J. E. Molecular architecture of the rotary motor in ATP synthase. *Science* **286**, 1700–1705 (1999).
24. Stroebel, D., Choquet, Y., Popot, J. L. & Picot, D. An atypical haem in the cytochrome b₆(f) complex. *Nature* **426**, 413–418 (2003).
25. Xia, D. et al. Crystal structure of the cytochrome bc₁ complex from bovine heart mitochondria. *Science* **277**, 60–66 (1997).
26. Zouni, A. et al. Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 Å resolution. *Nature* **409**, 739–743 (2001).
27. Binda, C., Newton-Vinson, P., Hubalek, F., Edmondson, D. E. & Mattevi, A. Structure of human monoamine oxidase B, a drug target for the treatment of neurological disorders. *Nature Struct. Biol.* **9**, 22–26 (2002).
28. Bracey, M. H., Hanson, M. A., Masuda, K. R., Stevens, R. C. & Cravatt, B. F. Structural adaptations in a membrane enzyme that terminates endocannabinoid signaling. *Science* **298**, 1793–1796 (2002).
29. Ferguson, K. M. et al. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* **11**, 507–517 (2003).
30. Newton, A. C. Regulation of the ABC kinases by phosphorylation: protein kinase C as a paradigm. *Biochem. J.* **370**, 361–371 (2003).
31. Kusumi, A. et al. Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 351–378 (2005).
32. Saxton, M. J. & Jacobson, K. Single-particle tracking: applications to membrane dynamics. *Annu. Rev. Biophys. Biomol. Struct.* **26**, 373–399 (1997).
33. Hessa, T. et al. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377–381 (2005).
34. Picot, D., Loll, P. J. & Garavito, R. M. The X-ray structure of the membrane protein prostaglandin H₂ synthase-1. *Nature* **367**, 243–249 (1994).

Acknowledgements I thank the members of my laboratory for discussions, and the US National Institutes of Health and NRCR for support.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to D.M.E. (don@mail.csbyale.edu).

1. Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science* **175**, 720–731 (1972).
2. Liu, Y., Engelman, D. M. & Gerstein, M. Genomic analysis of membrane protein families: abundance and conserved motifs. *Genome Biol.* **3**, 0054.1–0054.12 (2002).
3. Doura, A. K., Kobus, F. J., Dubrovsky, L., Hibbard, E. & Fleming, K. G. Sequence context

Solving the membrane protein folding problem

James U. Bowie¹

One of the great challenges for molecular biologists is to learn how a protein sequence defines its three-dimensional structure. For many years, the problem was even more difficult for membrane proteins because so little was known about what they looked like. The situation has improved markedly in recent years, and we now know over 90 unique structures. Our enhanced view of the structure universe, combined with an increasingly quantitative understanding of fold determination, engenders optimism that a solution to the folding problem for membrane proteins can be achieved.

In Ilse Aichinger's famous short story *The Bound Man*, a man awakes from a coma to find himself bound by ropes¹. He learns to move gracefully within these constraints and eventually becomes a circus performer. Similarly, membrane proteins must perform complex signalling and transport functions within the strict confines of a lipid bilayer. This requires elegant structural adaptations to their environment.

Our early views of membrane-protein structure were largely shaped by the pioneering work of Henderson and colleagues, who established a modest resolution view of bacteriorhodopsin in 1975 (ref. 2). Confirming the prescient prediction of Lenard and Singer in 1966 (ref. 3), the structure revealed a bundle of long helical rods traversing the membrane. This suggested that membrane proteins could be largely thought of as an assembly of transmembrane (TM) helices. This view was reinforced by the stretches of hydrophobic residues found in membrane-protein sequences that were long enough to form membrane-spanning helices⁴. But recent structures have challenged this simple concept of membrane-protein architecture. In this review, I will briefly illustrate some of the surprising twists and turns that polypeptide chains make in the bilayer and then discuss our current understanding how these structures are built. Finally, I will argue that with concerted effort, practical solutions to the membrane-protein-folding problem are possible.

Complex architectures

Protein structure within the bilayer can be divided into two general types: β -barrels and bundles of α -helices. Because the folding problem for β -barrels and helix bundles is very different, and because β -barrel proteins are much less common⁵, I will focus on the helix bundle class. Figure 1 shows our progress with helix bundle protein-structure determination since the first high-resolution membrane protein structure was solved 20 years ago (reviewed in ref. 6). A decade later, the pace of structure determination quickened. We now know 52 unique helix bundle structures, and 38 of these were solved in the past 5 years. As a result, we have a much clearer view of the structural diversity exhibited by membrane proteins. Although long TM helices are still considered a fundamental building block, polypeptide chains can be organized into other complex patterns.

Perhaps the first dramatic departure from the standard view was the structure of the glycerol/water channel GlpF, determined in 2000

(ref. 7). This structure illustrates some of the complexities of membrane-protein architecture. Two views of the GlpF monomer (the protein is a tetramer) are shown in Fig. 2, and I will call them front and back. The back side presents a simple picture of TM helices packed together roughly parallel to the membrane normal. Few would have been surprised by this image, even 30 years ago. The front side presents a more complex view, however. Most notable is the pair of helices that penetrate half way into the bilayer, where their amino-termini meet. Such half TM helices are not very common, but are also not particularly unusual (about 1 in 20 of all TM helices)⁸. Also visible in the front view is a common feature of membrane proteins: a highly distorted TM helix. About 60% of TM helices contain significant bends or other distortions⁹. Finally, loop conformations between

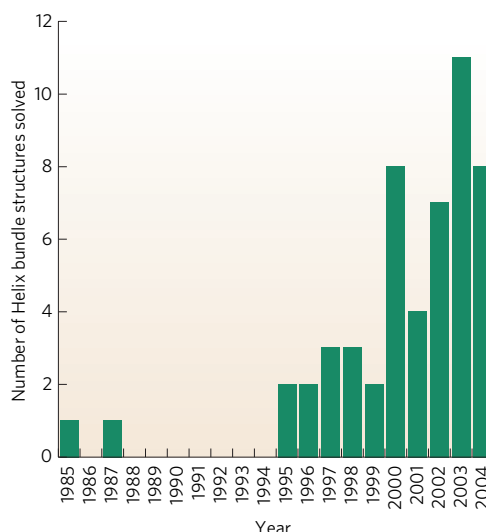


Figure 1 | Progress of helix bundle membrane protein structure determination. Only unique structures are included. The data were obtained from a website maintained by Stephen White (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html). For a similar plot including all membrane protein structures, see ref. 6.

¹Department of Chemistry and Biochemistry, UCLA-DOE Center for Genomics and Proteomics, Molecular Biology Institute, Boyer Hall, UCLA, 611 Charles E. Young Drive E, Los Angeles, California 90095-1570 USA.

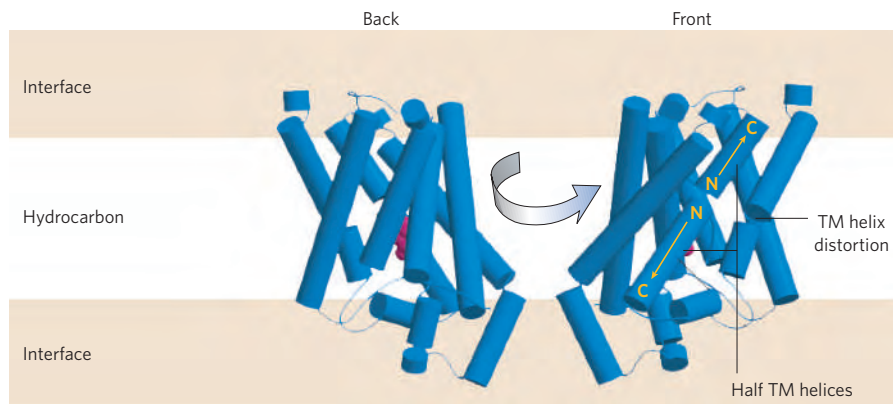


Figure 2 | The structure of the glycerol channel GlpF. Two views, labelled front and back, are shown⁷. They are rotated 180° with respect to each other along the membrane normal. The N and C termini of the half TM helices are labelled in the front view. A glycerol in the channel is shown in CPK mode. The structure was approximately positioned in the hydrocarbon core of the membrane by finding the most hydrophobic 30-Å slab of the protein perpendicular to the four-fold axis of the tetramer.

TM helices do not seem notably restricted by the membrane. Although packed TM helices remain by far the most common membrane-protein structural feature, protein chains are clearly not bound by such rigid limitations. Membrane proteins have developed tricks to get their chains to go where needed to satisfy functional imperatives. To solve the membrane-protein-folding problem we will need to understand the genesis of these structural idiosyncrasies.

Two fundamental stages in membrane-protein folding

As proposed originally by Popot and Engelman¹⁰, it is convenient to break membrane-protein folding into two stages: insertion and folding. These stages are illustrated in Fig. 3.

In the first stage, the membrane protein is inserted across the bilayer. This stage can be both directed and catalysed by a translocon complex. In this manner, some of the membrane-inserted segments, and their topology, are established. The first phase marks membrane protein folding as fundamentally different from soluble protein folding because the insertion and topology decisions are made in consultation with the translocon complex^{11,12}. By contrast, the fold of soluble proteins is normally defined entirely within the sequence itself¹³. To understand the first stage, we will need to learn how the sequence and the translocon communicate.

In the second stage, the tertiary and quaternary structures are built. This second stage involves assembly and reorientation of the TM segments established in the first phase, the additional insertion of re-entrant portions of the chain and subunit oligomerization. It may make sense to separate TM helix packing from other folding steps¹⁴, but I lump these interrelated steps together in this review. A complete description of the second stage will require an understanding of amino-acid preferences for different portions of the bilayer, the energetics of interactions within the protein itself as well as interactions between the protein and the bilayer.

A thermodynamic view of insertion

Unlike soluble proteins, membrane proteins reside in a variable and anisotropic environment¹. Figure 4 shows a snapshot from a molecular dynamics simulation of a fluid POPC (1-palmitoyl 2-oleoyl phosphatidylcholine) bilayer¹⁶. It highlights the conformational heterogeneity of the lipid conformations and variable molecular compositions in different regions¹⁷. Although somewhat arbitrary, the membrane is typically divided into two general regions. The hydrocarbon core in the centre is roughly 30 Å wide in a typical bilayer and is dominated by the aliphatic lipid chains. The interfacial region of the bilayer comprises lipid headgroups and considerable bound water. The interfacial region is quite large, roughly 15 Å across, and is therefore a major feature of a membrane protein's environment. The atom distributions shown in Fig. 4 illustrate how the properties of the bilayer change markedly along the direction of the bilayer normal. Starting from the centre, the environment changes from extremely apolar to highly polar/charged and finally to bulk water, all in a space of about 30 Å. Clearly, our descriptions of membrane-

protein-folding energetics will need to vary as a function of bilayer depth.

A characteristic feature of helical membrane-protein sequences is their hydrophobic stretches, which are approximately 20 residues long¹⁰. A 20-residue segment is just long enough to span the hydrocarbon core of a typical bilayer in a helical conformation. Hydrophobicity is an important feature defining TM helices, which implies that thermodynamic partitioning between the water and the bilayer plays a significant role in membrane insertion and/or maintenance of the membrane protein in the bilayer. Figure 5 shows free energies for transferring each of the 20 amino acids from water into octanol, which serves as a model for the hydrocarbon core^{15,18}. According to this scale, transferring the backbone (a glycine residue) is unfavourable by 1.25 kcal mol⁻¹ per residue, arguing that side-chain hydrophobicity drives the equilibrium in favour of insertion. In other words, there must be a threshold of side-chain hydrophobicity for the segment to favour bilayer insertion¹⁹. For example, according to this scale, a 20-residue polyalanine sequence would not be sufficiently hydrophobic to partition into the hydrocarbon core ($\Delta G = 10$ kcal mol⁻¹), but replacing five

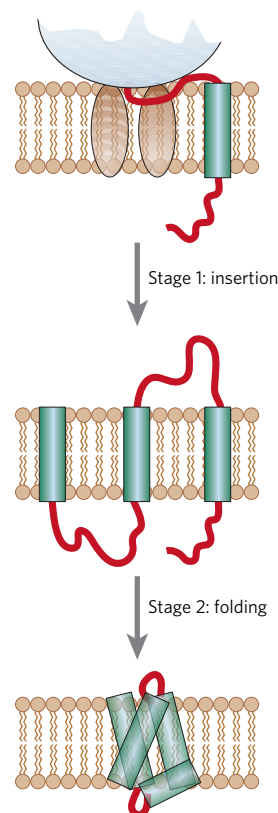


Figure 3 | Two stages of membrane protein folding¹⁰. See the text.

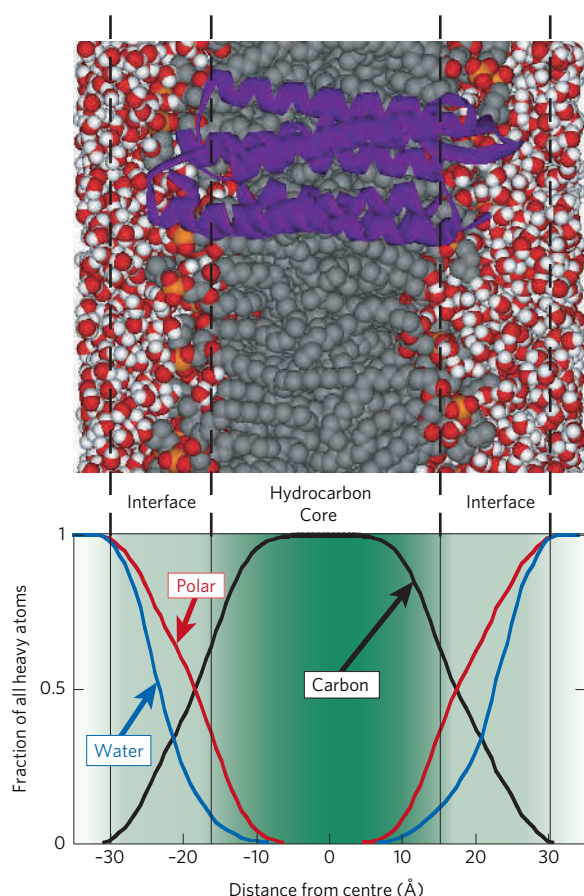


Figure 4 | A snapshot of a molecular dynamics simulation of a POPC bilayer¹⁶. The structure of bacteriorhodopsin is shown in front of the bilayer to give a sense of scale and was not included in the simulation. The graph at the bottom shows the fraction of carbon atoms (black line) the fraction of phosphorus and oxygen atoms generating polar moieties (red line) and the fraction of oxygen atoms from water (blue line). This figure was inspired by an earlier review by White and Wimley⁸⁹.

of the alanines with five leucines would be just favourable ($\Delta G = -1.25$ kcal mol⁻¹). Recent experiments on translocon-catalysed TM insertion indicate that this prediction is remarkably accurate²⁰. Nevertheless, octanol is not a perfect model of the hydrocarbon core, which is not a homogeneous solvent, and transfer free energies vary as a function of bilayer depth²¹. As discussed below, new translocon-insertion experiments promise significant advancements over this basic thermodynamic model.

A biological view of insertion

Many membrane proteins are inserted as they emerge from the ribosome through a protein-conducting channel called the SecY complex in bacteria and the Sec61 complex in eukaryotes. The SecY/Sec61 channel must do a number of extraordinary things as a nascent polypeptide emerges from the ribosome. First, it needs to allow passage of the polypeptide chain through the membrane without leakage of ions or other small molecules. Second, it needs to decide whether to reverse the orientation of the emerging segment of the protein chain. Third, it needs to decide whether the emerging segment of the protein should be passed through to the aqueous compartment on the other side or slid into the membrane. Fourth, if it decides that the segment should go into the membrane, it needs to open the channel laterally and pass the segment into the bilayer, again without membrane leakage. A structure of the SecY/Sec61 translocon complex from *Methanococcus jannaschii* has now been determined²², which markedly advances our ideas about how these complex functions can be accomplished.

The solved structure of the translocon complex, shown in Fig. 6, contains three polypeptide chains α , β and γ . The core of the complex is the large α subunit with ten TM helices surrounding the probable translocation pore. The structure is completed by the γ subunit, which closes off the α subunit on one end, and by the β subunit, which contributes an additional TM helix. The overall shape looking down from above or below the membrane resembles a clamshell that could possibly open on the side opposite the γ subunit (see Fig. 6). From the side, the channel is roughly hour-glass shaped and is accessible to water, but the aqueous channel is blocked by a short, helical plug from the α subunit. The structure solved is of the closed form of the channel without a translocating polypeptide. Nevertheless, the structure, coupled with considerable biochemical data, enabled Rapoport and colleagues^{22,23} to suggest a plausible model for the general protein translocation process, which I will briefly summarize.

As the protein emerges from the ribosome, the plug slides out of the way, allowing the polypeptide to pass through the channel. The constriction in the channel is lined by hydrophobic residues (dark blue ring), which may provide a flexible self-sealing mechanism, maintaining a barrier between the two compartments (other mechanisms may operate as well). If the segment of the protein is destined for the external aqueous compartment, this *status quo* can be maintained. If the segment needs to be inserted into the membrane, however, something else needs to happen. The present model suggests that TM helices can be passed into the membrane by opening the channel to the side opposite the γ subunit (see Fig. 6).

The topology decision

How is the topology of the membrane-inserted segments established? Gunnar von Heijne and colleagues discovered that the cytoplasmic side of membrane proteins tends to be positively charged: the so-called positive-inside rule²⁴. It is possible that part of this preference results from electrostatic potential differences between the two compartments and/or lipid composition differences between the two membrane leaflets. It is now clear, however, that the translocon itself plays an important role in the topology decision. Goder *et al.*¹² showed that by reversing the charge on positively and negatively charged residues in the *Saccharomyces cerevisiae* Sec61 α subunit, the topology of membrane proteins could be reversed. Thus, the same sequence can have a differ-

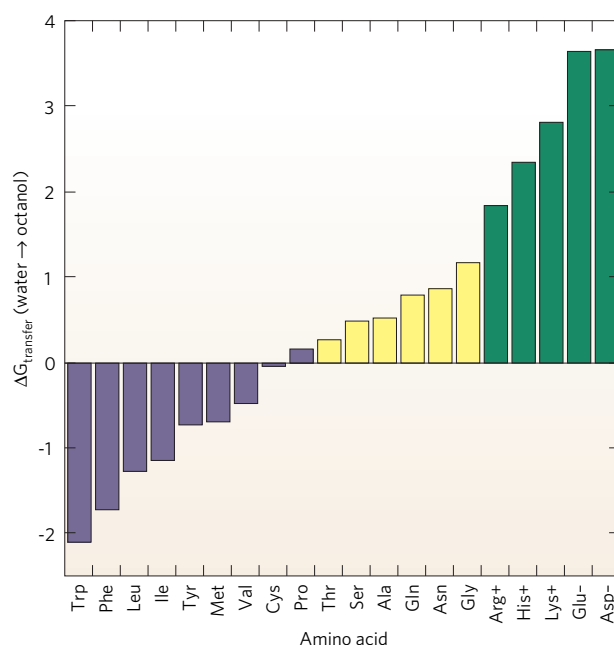


Figure 5 | Free energies for transfer of amino acids from water to octanol¹⁵. Charged residues are shown in green bars, polar residues in yellow bars and hydrophobic residues in purple bars.

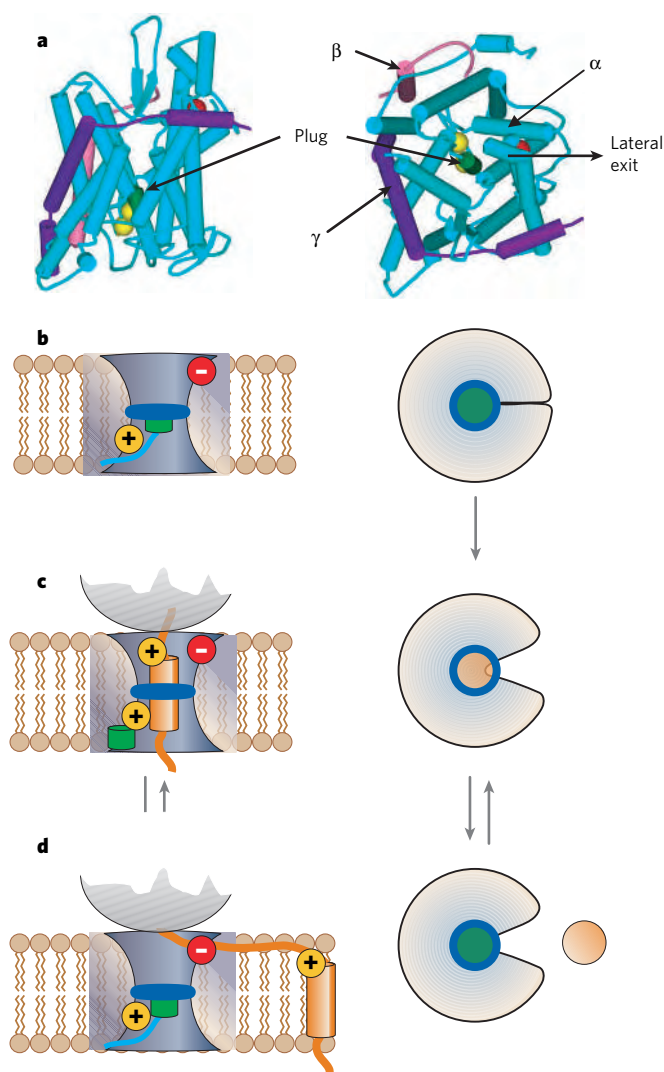


Figure 6 | The Sec61 translocon structure and mechanism model²³. **a**, The α subunit is shown in light blue, the β subunit in purple and the γ subunit in pink. The plug helix that blocks the pore is shown in green. Key charged residues that help define the topology of translocating polypeptides in *S. cerevisiae* are shown in yellow (positive charge) and red (negative charge). **b**, Schematic side and top views of the Sec61 translocon showing the locations of the charges, the plug helix in green and the hydrophobic collar shown by the dark blue ring. **c**, A schematic view of a translocon with a nascent polypeptide (orange) emerging from the ribosome (grey). A positively charged residue helps define the topology of this segment as N-terminal first. **d**, A schematic view of the translocon opening the lateral gate so the helix can exit to the membrane if the partitioning is favourable.

ent topology depending on the sequence of the translocon. The positions of the key charged residues identified in yeast can be mapped onto the *M. jannashii* structure as shown in Fig. 6 (refs 11, 23). Because of these charged residues, as the emerging polypeptide chain approaches the region of the plug it would probably experience a strong positive electrostatic potential. If the emerging segment was positively charged, electrostatic repulsion could drive a reorientation of the polypeptide, pushing the N terminus away from the constriction and toward the negative charge at the top of the pore, thereby flipping the orientation of the chain. Other factors also contribute to the topology decision, such as synthesis rate and the length of the hydrophobic segment^{12,23}. Moreover, the important charged residues identified in yeast are not conserved in all Sec61 translocons, suggesting either that alternative residues contribute to this mechanism or that other solutions to this problem operate in different organisms. This could be one of the reasons that it is so difficult to express membrane proteins in heterologous organisms.

The insertion decision

How does the translocon know which segments should be inserted into the membrane? Recent work by Hessa *et al.* has significantly improved our understanding of the insertion code²⁰. They examined the probability that a 19-residue segment would be inserted into the endoplasmic reticulum membrane or passed through the translocon and into the aqueous lumen. As discussed above, physical chemistry suggests that a polyalanine sequence would not be sufficiently hydrophobic to insert into the membrane, but if five alanines were replaced with leucine, it would be just at the threshold for insertion. Remarkably, this is exactly what was observed. As the hydrophobicity of a polyalanine segment was increased by adding leucines, the insertion probability increased. Moreover, the insertion-probability increase could well be described by a Boltzmann distribution, as though the segments partition according to their transfer free energies. By substituting different amino acids into the test segment, it was possible to measure an apparent free energy of biological partition for each of the 20 amino acids. The biological scale correlates reasonably well with measured free energies for transferring amino-acid side chains from water to octanol (see above). Thus, the insertion probability behaves as though the translocon is measuring the free energy of partitioning between the aqueous and membrane phases.

How can the translocon detect the partitioning free energy? A plausible model is that the translocon frequently opens the lateral gate during protein translocation (see Fig. 6)²⁵. When this happens, the nascent polypeptide segment in the translocon channel can sample both the aqueous and membrane phases. If the rate of synthesis is sufficiently slow relative to this sampling process, equilibrium can be established. If the equilibrium favours the bilayer, the segment becomes inserted. Otherwise it continues straight through the channel. In this manner, the translocon can act as a catalyst for the partitioning process. The structural details of this opening and how it can be accomplished without leakage is still something of a mystery.

The idea that the translocon catalyses partitioning between the aqueous and membrane phases creates a wonderful opportunity for refining our understanding of insertion thermodynamics beyond the model solvent studies discussed above. Because the membrane is not a homogeneous environment, insertion probability does not depend just on composition, but also on the location of individual residues within the TM segment. Insertion probabilities for polar residues are particularly dependent on their position in the TM segment²⁰. These differences could be due to variations in the environment at different bilayer depths, as well as to compensating structural changes in both the bilayer and the protein. For example, long polar amino acids can arrange their side chains to move their polar atoms away from the hydrocarbon core in the direction of the increasing polarity gradient. This so-called 'snorkelling'²⁶ can make the placement of polar amino acids increasingly favourable as they approach the interfacial region. Translocon-insertion experiments suggest that it is about 2 kcal mol⁻¹ more favourable to place an arginine residue near the edge of the hydrocarbon core than at the centre²⁷. This factor is so significant that it is even possible to insert the S4 helix of a voltage-gated potassium channel, which contains four arginine residues, into a membrane²⁷. Moving beyond simple consideration of segment hydrophobicity to define factors that contribute to insertion probability will greatly improve our ability to predict which segments of a protein become inserted. This is a key step on the way to solving the membrane-protein-folding problem.

Post-insertion folding

After (and probably during) insertion, the protein can begin a search for the stable, native conformation. We currently have little experimental knowledge of what the membrane-inserted unfolded state looks like or the pathway to the folded state in bilayers. Nevertheless, it is reasonable to surmise that a major component of the native-state conformational search will be packing of the stably inserted helices

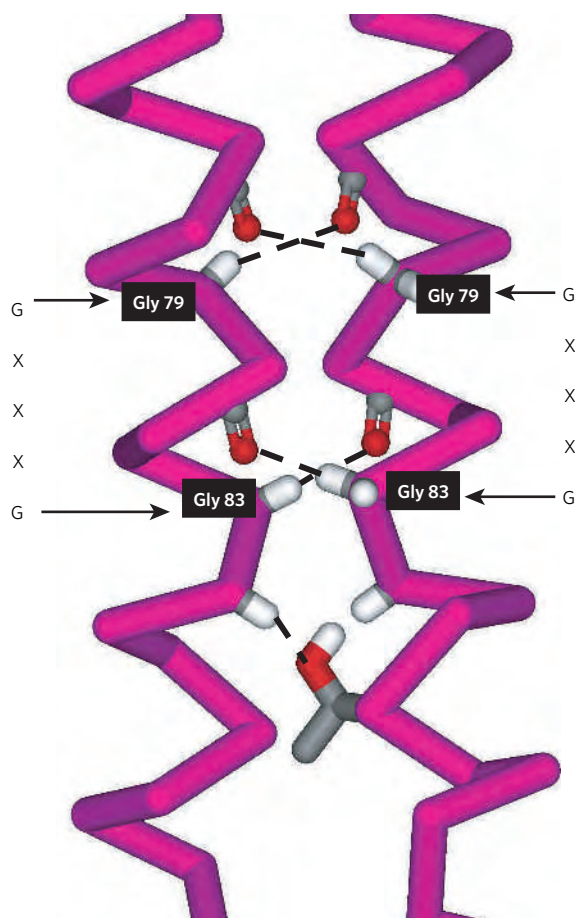


Figure 7 | The glycoporphin A dimer⁸⁹. The two key packing interface glycines found in the GXXXG sequence are highlighted. A network of C α -H...O hydrogen bonds between backbone carbonyls and the C α hydrogens of the glycine residues are shown, along with a C α -H...O hydrogen bond made by a threonine side chain. In the atoms shown, carbon is grey, oxygen is red and hydrogen is white. They hydrogen bonds are indicated by the dashed lines.

from stage 1 (refs 10, 14). Considerably more needs to happen, however, to attain a complex structure such as GlpF (Fig. 2). Additional segments of the protein chain must be inserted into the membrane or at least rearranged within the membrane. We also cannot rule out the possibility of topology rearrangements during folding. Some membrane proteins can spontaneously insert into bilayers²⁸, and Dowhan and co-workers have found that the topology of improperly inserted membrane proteins can self-correct if the lipid composition is altered after incorrect insertion, indicating that hydrophilic loop regions can be dragged through the bilayer²⁹.

Like the soluble-protein-folding problem, folding of membrane proteins probably proceeds down a funnel-shaped energy landscape to an energy minimum³⁰. Consistent with a folding funnel view is the observation of multiple pathways in the folding of bacteriorhodopsin³¹. A major difference between soluble protein folding and membrane protein folding, however, is that the starting point is much more constrained, because much of the secondary structure and topology will be set by the insertion process. Thus, the unfolded protein is much farther down the folding funnel and closer to the folded state compared with soluble proteins. It seems clear that the folding-energy landscape is defined by a complex interplay between various forces, including polypeptide partitioning in the bilayer (discussed above), interactions between lipid and protein and interactions within the protein itself. I will briefly discuss these factors in the following sections. I will focus on forces acting within the bilayer because the extensive work on water-soluble protein folding will largely be applicable to chain folding in the water-exposed portions.

Basic features of transmembrane helix packing

As a consequence of environmental constraints, TM helices exhibit a narrower distribution of packing angles than found in soluble proteins, with a strong preference at around 20° (ref. 32). In fact, more than 60% of all TM-helix-packing angles fall in the range of 0° to 40°. Part of this preference is due to the favourable orientation of TM helices along the membrane normal, which tends to favour small packing angles³³. The 20° packing angle is particularly favourable because it facilitates inter-helix side-chain interdigitation³⁴. TM helices can also exhibit super-helical wrapping³⁵, which limits the helix divergence that would occur if they remained perfectly straight.

TM-helix-packing interfaces show a preference for small residues³⁶. A well-known example is the strong TM helix dimer from glycoporphin A shown in Fig. 7. The core of this interface, a GXXXG sequence motif, is found in many helix oligomers³⁷. The glycine residues in this motif face each other in the dimer, allowing close packing of the helices. There are several possible explanations for the small-residue preference. First, small residues allow close approach of the helices, leading to improved packing³⁸. Second, there is a lower entropy cost for fixing the conformation of small side chains in the folded protein compared with larger residues³⁹. Finally, small side chains expose polar backbone atoms that could lead to favourable interactions^{40,41} (see below). As a measure of the significance of this preference, a simple scoring function that considers only residue size is surprisingly effective for packing TM helices⁴².

Deviant helices

TM helices contain many local distortions including kinks and short stretches of π or 3_{10} helices, leading to a deviation of the helix axis and/or a shift in the regularity of side-chain positions^{43,44}. Given their prevalence, predicting the location of helix distortions will be an important piece of the protein-folding puzzle. There are now encouraging signs. For example, Yohannan *et al.* found that kinks could be identified with more than 90% reliability simply by looking at proline abundance in a multiple sequence alignment⁹. Rigoutsos *et al.* report the identification of sequence patterns predictive of helix distortions⁴⁵, and local sequence features have also been suggested to alter kink magnitude⁴⁶.

Why are TM helices so frequently kinked? One probable reason is that kinks enable the small structural adjustments needed to position functional groups precisely, which could facilitate functional diversification of a common architecture⁹. For example, G-protein-coupled receptors all have seven TM helices but respond to a remarkable range of signals. This would probably be impossible if their construction were limited to seven rigid rods. Another possibility is that the helical distortions provide weak points in the helical rods that facilitate movements needed for protein function⁴⁴. Finally, proline residues were found to block off-pathway β -sheet formation during CFTR (cystic fibrosis transmembrane conductance regulator) folding⁴⁷. Thus, some proline residues can be important for directing folding toward the native state structure.

Internal forces stabilizing membrane proteins

TM regions are dominated by apolar side chains, and van der Waals interactions play an important role in stabilizing interactions in the hydrocarbon core region of the bilayer. For example, the interface of the synaptobrevin helix dimer seems to be constructed entirely from apolar side chains⁴⁸. On the basis of the stability effects of side-chain deletions in bacteriorhodopsin and glycoporphin A, side-chain burial contributes approximately 27 cal mol⁻¹ Å⁻² (in detergent), which is similar to the contribution observed in soluble proteins⁴⁹. In soluble proteins, a significant fraction of the energetic contribution of residue burial arises from the hydrophobic effect, which is not a dominant factor in the hydrocarbon core region of the bilayer. Somehow membrane proteins seem to make up for the loss of the hydrophobic effect. One possible factor is that packing is better in membrane proteins than soluble proteins, but there is no consensus on this point. Although by some measures membrane proteins are better packed than in soluble

proteins^{50,51}, others argue that membrane proteins contain more packing defects^{52,53}. It is possible that locally well-packed regions of structure are particularly important and more than compensate for the packing defects observed elsewhere. Alternatively, the unfolded state of membrane proteins may pack less well with the bilayer than soluble proteins pack with water. Although the packing density of water is lower (0.36; ref. 54) than lipids in bilayers (about 0.42 in the centre and about 0.65 near headgroups)⁵⁵, water is freer to pack around the protein side chains than are lipids.

Although apolar residues dominate in bilayer-embedded regions of membrane proteins, the interiors of membrane proteins are more polar than their exteriors⁵⁶ and there is ample potential for hydrogen-bonding interactions⁵⁷. In aqueous solution, the strength of hydrogen bonds is diminished by the high dielectric effect in water and competition from water for hydrogen bonds. In the apolar hydrocarbon core, however, hydrogen bonds have the potential to be very strong^{15,58}, worth possibly more than 5 kcal mol⁻¹ (ref. 15). Experimental measurements of hydrogen-bond contributions between TM helices are significantly lower than that, however, and quite variable with an upper limit in the range of 2 kcal mol⁻¹ (refs 49, 59).

Several factors can reduce hydrogen-bond contributions. First, the position in the bilayer seems to be very important⁵⁹. If placed near the edge of the hydrocarbon core, polar moieties from the interfacial region can make competing hydrogen bonds in the unfolded state, reducing the net contribution to stability⁵⁹. Second, it may be difficult to arrange good hydrogen-bond geometries given the constraints of the polypeptide chain in a folded structure^{49,60}. Nevertheless, appropriately placed hydrogen bonds with good geometry can be quite significant. For example, the introduction of a single polar residue into an apolar TM helix can strongly drive oligomerization^{59,60-64}. Networks of polar interactions have been described that probably play important roles in defining TM helix structure⁵⁷.

Because of the potential for forming strong hydrogen-bonding interactions, polar residues must be used judiciously in membrane proteins or they could engage in inappropriate interactions, particularly in the crowded membrane environment⁶³. Indeed, polar substitutions are the most common disease-causing mutations in membrane proteins⁶⁵. For example, one of the earliest identified mutations in an oncogene was the valine to glutamic acid mutation in the TM domain of the Her2/Neu receptor. The glutamic acid forms a hydrogen bond in a Neu TM domain peptide⁶⁶ and the mutation seems to induce receptor aggregation and consequent activation⁶⁷. In the CFTR chloride channel, a valine to asparagine mutation also seems to alter TM helix interactions through inappropriate hydrogen bonding⁶⁸.

In addition to strongly polar interactions, weakly polar interactions probably play a significant role in stabilizing membrane-protein structure⁵⁸. Of particular note are C α -H...O hydrogen bonds⁴⁰. Interest in these carbon-hydrogen bonds was piqued by *ab initio* quantum mechanics calculations suggesting that C α -H...O hydrogen bonds could be approximately half the strength of a traditional hydrogen bond⁶⁹. Senes *et al.* demonstrated that they are also prevalent in membrane-protein structures⁴⁰. An example of a striking network of potential C α -H...O hydrogen bonds is found in the strong TM helix dimer of glycophorin A, shown in Fig. 7, which is made possible by the small glycine residues in the interface noted above.

So how strong are these carbon-hydrogen bonds? Like traditional hydrogen bonds, it seems that the answer depends on the context. In an elegant test, Arbely and Arkin found that hydrogen bonding altered the C α -H bond stretching frequency and estimated a contribution of 0.9 kcal mol⁻¹ (ref. 41). This may be an overestimate as it is based on *in vacuo* model data. Nevertheless, these results indicate that the interaction is favourable. By contrast, Yohannan *et al.* found that a C α -H...O hydrogen bond in bacteriorhodopsin made little contribution to stability⁷⁰. Mottamal and Lazaridis suggest that the difference in these results stems from the different hydrogen-bond geometries⁷¹. Thus, like normal hydrogen bonds, the structural details will probably

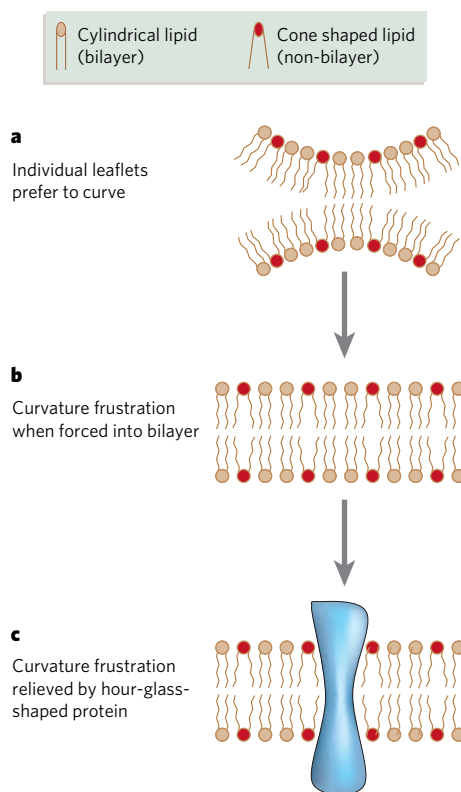


Figure 8 | Curvature elastic energy. Two types of lipid are depicted as either cylindrical or cone-shaped. **a**, The presence of a cone-shaped lipid in a bilayer will cause the two leaflets to want to curve away from one another. **b**, Forcing them into a bilayer causes overpacking in the hydrocarbon tails. **c**, An hour-glass-shaped protein can release some of this stored curvature elastic energy.

define C α -H...O hydrogen-bond strength. Because of the importance of hydrogen bonding in the membrane interior, a detailed understanding of factors that alter the strengths of hydrogen bonds will be an important step toward a solution to the membrane-protein-folding problem.

Non-specific driving forces in the bilayer

In addition to polarity gradients affecting side-chain partitioning discussed earlier, other bilayer properties can play a major role in membrane-protein folding, stability and structure⁷². Properties of specific lipids are treated in detail in an accompanying review. Here we will simply consider general effects. Although it is more correct to describe lipid packing in terms of intermolecular forces⁷², in qualitative terms it is useful to think of different lipids as having distinct headgroup and hydrocarbon tail sizes (see Fig. 8). To form a bilayer, the lipid should have a cylindrical shape with the radius of the headgroup approximately matching the radius of the hydrocarbon tails. In some lipids, the hydrocarbon tails splay out so they cannot form a bilayer by themselves. Such 'non-bilayer' lipids can be incorporated into a bilayer if they are mixed with bilayer-forming lipids. As shown in Fig. 8, the shape mismatch creates a driving force for the bilayer leaflets to curve away from one another. Forcing them into a bilayer creates a curvature frustration⁷³, which can be simplistically described as an overpacking in the hydrocarbon region and an underpacking in the headgroup region.

The drive to relieve curvature elastic energy can shape membrane protein structure. For example, as shown in Fig. 8, an hour-glass shape should relieve the curvature strain. Indeed, Tamm and co-workers showed that curvature elastic energy has a significant effect on the stability of OmpA, a hour-glass-shaped β -barrel protein⁷⁴. Also, Booth and co-workers have demonstrated that changes in lipid composition

predicted to change lateral packing pressure alter the folding pathway of bacteriorhodopsin⁷⁵. The lack of curvature elastic energy in detergent micelles could be a factor in destabilizing or altering the structure of membrane proteins in detergent solution.

Different lipid compositions can alter the natural bilayer thickness. If the hydrophobic surface of the membrane protein is thicker or thinner than the hydrocarbon core of the bilayer, a tension is created because of the drive to shield hydrophobic surfaces from water. As shown in Fig. 9, this hydrophobic mismatch tension can be resolved in several ways. Either the protein can adjust to match the bilayer or the bilayer can deform to match the protein. *In vitro* experiments in purified systems demonstrate that in β -barrel proteins, the lipid deforms to match the protein^{74,76}. By contrast, helical membrane proteins, which may be less rigid than β -barrel proteins, seem able to adjust to the hydrophobic thickness of the bilayer^{77,78}. However, recent experiments with membranes isolated from cells that are filled with helical proteins suggest that membrane thickness is set by the proteins, not the lipids⁷⁹. One possible explanation for this difference is the complex lipid composition in natural membranes that could facilitate the bilayer deformations necessary for hydrophobic matching (see Fig. 9). Indeed, curvature elastic strain does seem to facilitate hydrophobic matching of the β -barrel protein OmpA⁷⁴.

Specific lipid interactions

In addition to general bilayer properties, bound lipids have been seen in many crystal structures⁷² and are often crucial for protein function. For example, cytochrome C oxidase is inactivated by the removal of cardiolipin⁸⁰ and KcsA requires anionic phospholipids⁸¹. Thus, lipids can act as cofactors for some membrane proteins and stabilize their structures.

Toward a solution to the membrane-protein-folding problem

A solution to the membrane protein-folding problem will probably follow the two stages outlined at the beginning of this review. We must

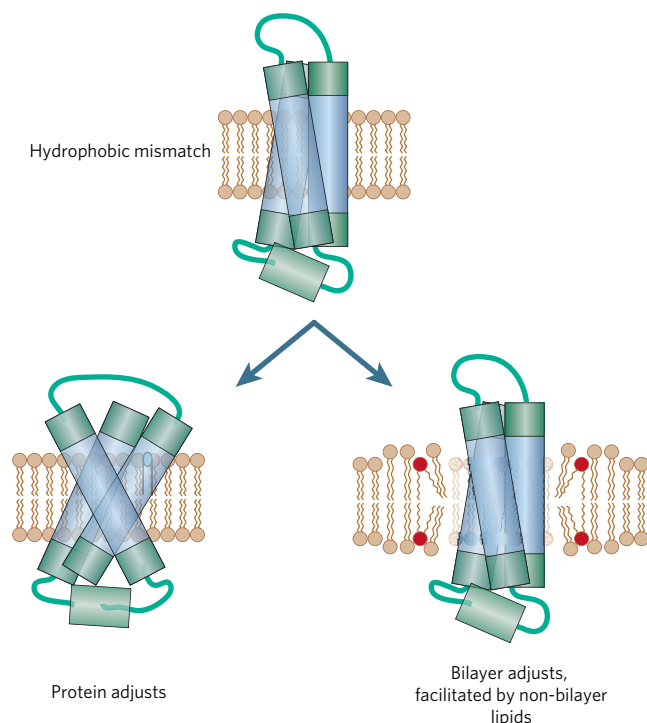


Figure 9 | Hydrophobic mismatch. If the hydrophobic region (blue) of a protein is thicker than the bilayer hydrocarbon core, either the protein can thin or the bilayer can thicken. Bilayer adjustment can be facilitated by non-bilayer lipids.

first be able to accurately predict the location and topology of the initially inserted segments and then learn to fold the chain from this starting point. It seems likely that we will first have a practical solution for stage 1. Although much work remains, TM-helix prediction methods are already quite accurate at finding TM helices, if not their endpoints⁸. New experiments allowing us to effectively query the insertion code²⁰ along with the increasing structure database will improve these methods further. Moreover, perfection in first-stage prediction may not be necessary because it should be possible to start second-stage folding with a collection of topology models.

Can we solve the second-stage folding problem? In considering this question, it is useful to make a comparison with soluble protein folding. Although there has been limited practical success in *de novo* soluble protein folding, the best results are obtained with low-contact-order (many interactions close in sequence) helical proteins, that is, those that are similar to helix bundle membrane proteins⁸². With membrane proteins, we can add significant additional constraints to folding algorithms, however. In particular, we can start from a first-stage model that is closer to the final fold than the starting point for soluble protein folding. Moreover, the bilayer can be used to great advantage for structure-prediction efforts because it provides spatial information. In membrane proteins, second-stage folding can be directed not only by intra-protein interactions but also by bilayer position. For example, a trial conformation with a highly charged segment in the middle of the bilayer can be quickly rejected in the absence of mitigating factors. If soluble-protein folders could add this type of spatial information, it seems likely that robust prediction methods would be at hand. Thus, it is reasonable to think that reliable prediction methods for membrane proteins could be developed if the same level of effort were directed toward this goal as has been applied to soluble proteins. Indeed, practical structure prediction of TM-helix packing has already been possible in simple systems with the addition of limited experimental constraints^{48,83–87}. To fold proteins with complexities such as half TM helices we will need to move beyond TM-helix packing algorithms and into global folding algorithms that can capture all the forces that drive such structural oddities. Ultimately, this will require a sophisticated description of the complex balance of forces that operate at different bilayer depths.

Our understanding of membrane-protein-folding determinants is still rudimentary and our database of structures is tiny in comparison with that of soluble proteins. But like the audience in Ilse Aichinger's imagined *Bound Man* witnessing the amazing physical adaptations of the man bound by ropes, we are beginning to comprehend how proteins adjust to limitations of the membrane. As indicated in this review, we are developing an increasingly quantitative understanding of the energetic forces that operate in a bilayer, and the pace of structure determination is accelerating. With the conformational restrictions imposed on membrane proteins, we are already proceeding more rapidly to effective structure prediction than has been possible with soluble proteins. Given the technical difficulties in obtaining membrane-protein structures, this is an area where prediction methods would be particularly welcome.

1. Aichinger, I. in *The Art of the Tale* (ed. Halpern, D.) 9–17 (Penguin Books, New York, NY, 1956).
2. Henderson, R. & Unwin, P. N. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257**, 28–32 (1975).
3. Lenard, J. & Singer, S. Protein conformation in cell membrane preparations as studied by optical rotatory dispersion and circular dichroism. *Proc. Natl Acad. Sci. USA* **56**, 1828–1835 (1966).
4. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
5. Bigelow, H. R., Petrey, D. S., Liu, J., Przybylski, D. & Rost, B. Predicting transmembrane beta-barrels in proteomes. *Nucleic Acids Res.* **32**, 2566–2577 (2004).
6. White, S. H. The progress of membrane protein structure determination. *Protein Sci.* **13**, 1948–1949 (2004).
7. Fu, D. et al. Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481–486 (2000).

8. Cuthbertson, J. M., Doyle, D. A. & Sansom, M. S. Transmembrane helix prediction: a comparative evaluation and analysis. *Protein Eng. Des. Sel.* **18**, 295–308 (2005).
9. Yohannan, S., Faham, S., Yang, D., Whitelegge, J. P. & Bowie, J. U. The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc. Natl Acad. Sci. USA* **101**, 959–963 (2004).
10. Popot, J. & Engelman, D. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* **29**, 4031–4037 (1990).
11. White, S. H. & von Heijne, G. The machinery of membrane protein assembly. *Curr. Opin. Struct. Biol.* **14**, 397–404 (2004).
12. Goder, V., Junne, T. & Spiess, M. Sec61p contributes to signal sequence orientation according to the positive-inside rule. *Mol. Biol. Cell* **15**, 1470–1478 (2004).
13. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
14. Engelman, D. M. *et al.* Membrane protein folding: beyond the two stage model. *FEBS Lett.* **555**, 122–125 (2003).
15. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365 (1999).
16. Tieleman, D. P., Sansom, M. S. & Berendsen, H. J. Alamethicin helices in a bilayer and in solution: molecular dynamics simulations. *Biophys. J.* **76**, 40–49 (1999).
17. Wiener, M. C. & White, S. H. Structure of a fluid dioleoylphosphatidylcholine bilayer determined by joint refinement of X-ray and neutron diffraction data. III. Complete structure. *Biophys. J.* **61**, 437–447 (1992).
18. Wimley, W. C., Creamer, T. P. & White, S. H. Solvation energies of amino acid side chains and backbone in a family of host-guest pentapeptides. *Biochemistry* **35**, 5109–5124 (1996).
19. Liu, L. P., Li, S. C., Goto, N. K. & Deber, C. M. Threshold hydrophobicity dictates helical conformations of peptides in membrane environments. *Biopolymers* **39**, 465–470 (1996).
20. Hessa, T. *et al.* Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377–381 (2005).
21. Wimley, W. C. & White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Struct. Biol.* **3**, 842–848 (1996).
22. Van den Berg, B. *et al.* X-ray structure of a protein-conducting channel. *Nature* **427**, 36–44 (2004).
23. Rapoport, T. A., Goder, V., Heinrich, S. U. & Matlack, K. E. Membrane-protein integration and the role of the translocation channel. *Trends Cell Biol.* **14**, 568–575 (2004).
24. von Heijne, G. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the transmembrane topology. *EMBO J.* **5**, 3021–3027 (1986).
25. Heinrich, S. U., Mothes, W., Brunner, J. & Rapoport, T. A. The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain. *Cell* **102**, 233–244 (2000).
26. Chamberlain, A. K., Lee, Y., Kim, S. & Bowie, J. U. Snorkeling preferences foster an amino acid composition bias in transmembrane helices. *J. Mol. Biol.* **339**, 471–479 (2004).
27. Hessa, T., White, S. H. & von Heijne, G. Membrane insertion of a potassium-channel voltage sensor. *Science* **307**, 1427 (2005).
28. Nagy, J. K., Lonzer, W. L. & Sanders, C. R. Kinetic study of folding and misfolding of diacylglycerol kinase in model membranes. *Biochemistry* **40**, 8971–8980 (2001).
29. Zhang, W., Campbell, H. A., King, S. C. & Dowhan, W. Phospholipids as determinants of membrane protein topology: Phosphatidylethanolamine is required for the proper topological organization of the gamma-aminobutyric acid permease (GabP) of *Escherichia coli*. *J. Biol. Chem.* **280**, 26032–26038 (2005).
30. Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19 (1997).
31. Lu, H. & Booth, P. J. The final stages of folding of the membrane protein bacteriorhodopsin occur by kinetically indistinguishable parallel folding paths that are mediated by pH. *J. Mol. Biol.* **299**, 233–243 (2000).
32. Bowie, J. Helix packing in membrane proteins. *J. Mol. Biol.* **272**, 780–789 (1997).
33. Hushilt, J., Millman, B. & Davis, J. Orientation of alpha-helical peptides in a lipid bilayer. *Biochim. Biophys. Acta* **979**, 139–141 (1989).
34. Chothia, C., Levitt, M. & Richardson, D. Helix to helix packing in proteins. *J. Mol. Biol.* **145**, 215–250 (1981).
35. Langosch, D. & Heringa, J. Interaction of transmembrane helices by a knobs-into-holes packing characteristic of soluble coiled coils. *Proteins* **31**, 150–159 (1998).
36. Jiang, S. & Vakser, I. A. Side chains in transmembrane helices are shorter at helix-helix interfaces. *Proteins* **40**, 429–435 (2000).
37. Curran, A. R. & Engelman, D. M. Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr. Opin. Struct. Biol.* **13**, 412–417 (2003).
38. Jiang, S. & Vakser, I. A. Shorter side chains optimize helix-helix packing. *Protein Sci.* **13**, 1426–1429 (2004).
39. MacKenzie, K. & Engelman, D. Structure-based prediction of the stability of transmembrane helix-helix interactions: the sequence dependence of glycophorin A dimerization. *Proc. Natl Acad. Sci. USA* **95**, 3583–3590 (1998).
40. Senes, A., Ubarretxena-Belandia, I. & Engelman, D. M. The Cα–H–O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl Acad. Sci. USA* **98**, 9056–9061 (2001).
41. Arbely, E. & Arkin, I. T. Experimental measurement of the strength of a C_α–H–O bond in a lipid bilayer. *J. Am. Chem. Soc.* **126**, 5362–5363 (2004).
42. Fleishman, S. J. & Ben-Tal, N. A novel scoring function for predicting the conformations of tightly packed pairs of transmembrane alpha-helices. *J. Mol. Biol.* **321**, 363–378 (2002).
43. Riek, R. P., Rigoutsos, I., Novotny, J. & Graham, R. M. Non-alpha-helical elements modulate polytopic membrane protein architecture. *J. Mol. Biol.* **306**, 349–362 (2001).
44. Cordes, F. S., Bright, J. N. & Sansom, M. S. Proline-induced distortions of transmembrane helices. *J. Mol. Biol.* **323**, 951–960 (2002).
45. Rigoutsos, I., Riek, P., Graham, R. M. & Novotny, J. Structural details (kinks and non-alpha conformations) in transmembrane helices are intrahelically determined and can be predicted by sequence pattern descriptors. *Nucleic Acids Res.* **31**, 4625–4631 (2003).
46. Deupi, X. *et al.* Ser and Thr residues modulate the conformation of pro-kinked transmembrane alpha-helices. *Biophys. J.* **86**, 105–115 (2004).
47. Wigley, W. C. *et al.* A protein sequence that can encode native structure by disfavoring alternate conformations. *Nature Struct. Biol.* **9**, 381–388 (2002).
48. Fleming, K. G. & Engelman, D. M. Computation and mutagenesis suggest a right-handed structure for the synaptobrevin transmembrane dimer. *Proteins* **45**, 313–317 (2001).
49. Faham, S. *et al.* Side-chain contributions to membrane protein structure and stability. *J. Mol. Biol.* **335**, 297–305 (2004).
50. Eilers, M., Patel, A. B., Liu, W. & Smith, S. O. Comparison of helix interactions in membrane and soluble alpha-bundle proteins. *Biophys. J.* **82**, 2720–2736 (2002).
51. Eilers, M., Shekar, S. C., Shieh, T., Smith, S. O. & Fleming, P. J. Internal packing of helical membrane proteins. *Proc. Natl Acad. Sci. USA* **97**, 5796–5801 (2000).
52. Hildebrand, P. W., Rother, K., Goede, A., Preissner, R. & Frommel, C. Molecular packing and packing defects in helical membrane proteins. *Biophys. J.* **88**, 1970–1977 (2005).
53. Adamian, L. & Liang, J. Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J. Mol. Biol.* **311**, 891–907 (2001).
54. Richards, F. M. The interpretation of protein structures: total volume, group volume distributions and packing density. *J. Mol. Biol.* **82**, 1–14 (1974).
55. Marrink, S., Sok, R. & Berendsen, H. J. Free volume properties of a simulated lipid membrane. *J. Chem. Phys.* **104**, 9090–9099 (1996).
56. Rees, D., DeAntonio, L. & Eisenberg, D. Hydrophobic organization of membrane proteins. *Science* **245**, 510–513 (1989).
57. Adamian, L. & Liang, J. Interhelical hydrogen bonds and spatial motifs in membrane proteins: polar clamps and serine zippers. *Proteins* **47**, 209–218 (2002).
58. Chamberlain, A. K., Faham, S., Yohannan, S. & Bowie, J. U. Construction of helix-bundle membrane proteins. *Adv. Protein Chem.* **63**, 19–46 (2003).
59. Lear, J. D., Gratkowski, H., Adamian, L., Liang, J. & DeGrado, W. F. Position-dependence of stabilizing polar interactions of asparagine in transmembrane helical bundles. *Biochemistry* **42**, 6400–6407 (2003).
60. Zhou, F. X., Merianos, H. J., Brunger, A. T. & Engelman, D. M. Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl Acad. Sci. USA* **98**, 2250–2255 (2001).
61. Choma, C., Gratkowski, H., Lear, J. D. & DeGrado, W. F. Asparagine-mediated self-association of a model transmembrane helix. *Nature Struct. Biol.* **7**, 161–166 (2000).
62. Li, R. *et al.* Activation of integrin αIIbβ3 by modulation of transmembrane helix associations. *Science* **300**, 795–798 (2003).
63. Gratkowski, H., Lear, J. D. & DeGrado, W. F. Polar side chains drive the association of model transmembrane peptides. *Proc. Natl Acad. Sci. USA* **98**, 880–885 (2001).
64. Zhou, F. X., Cocco, M. J., Russ, W. P., Brunger, A. T. & Engelman, D. M. Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nature Struct. Biol.* **7**, 154–160 (2000).
65. Partridge, A. W., Therien, A. G. & Deber, C. M. Missense mutations in transmembrane domains of proteins: phenotypic propensity of polar residues for human disease. *Proteins* **54**, 648–656 (2004).
66. Smith, S., Smith, C. & Bormann, B. Strong hydrogen bonding interactions involving a buried glutamic acid in the transmembrane sequence of the neu/erbB-2 receptor. *Nature Struct. Biol.* **3**, 252–258 (1996).
67. Weiner, D., Liu, J., Cohen, J., Williams, W. V. & Green, M. A point mutation in the neu oncogene mimics ligand induction of receptor aggregation. *Nature* **339**, 230–231 (1989).
68. Therien, A. G., Grant, F. E. & Deber, C. M. Interhelical hydrogen bonds in the CFTR membrane domain. *Nature Struct. Biol.* **8**, 597–601 (2001).
69. Vargas, R., Garza, J., Dixon, D. A. & Hay, B. P. How strong is the Cα–H···O=C hydrogen bond? *J. Am. Chem. Soc.* **122**, 4750–4755 (2000).
70. Yohannan, S. *et al.* A Cα–H···O hydrogen bond in a membrane protein is not stabilizing. *J. Am. Chem. Soc.* **126**, 2284–2285 (2004).
71. Mottamal, M. & Lazaridis, T. The contribution of Cα–H···O hydrogen bonds to membrane protein stability depends on the position of the amide. *Biochemistry* **44**, 1607–1613 (2005).
72. Lee, A. G. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta* **1666**, 62–87 (2004).
73. Gruner, S. M. Intrinsic curvature hypothesis for biomembrane lipid composition: a role for nonbilayer lipids. *Proc. Natl Acad. Sci. USA* **82**, 3665–3669 (1985).
74. Hong, H. & Tamm, L. K. Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proc. Natl Acad. Sci. USA* **101**, 4065–4070 (2004).
75. Allen, S. J., Curran, A. R., Templer, R. H., Meijberg, W. & Booth, P. J. Controlling the folding efficiency of an integral membrane protein. *J. Mol. Biol.* **342**, 1293–1304 (2004).
76. O’Keeffe, A. H., East, J. M. & Lee, A. G. Selectivity in lipid binding to the bacterial outer membrane protein OmpF. *Biophys. J.* **79**, 2066–2074 (2000).
77. Williamson, I. M., Alvis, S. J., East, J. M. & Lee, A. G. Interactions of phospholipids with the potassium channel KcsA. *Biophys. J.* **83**, 2026–2038 (2002).
78. Powl, A. M., East, J. M. & Lee, A. G. Lipid-protein interactions studied by introduction of a tryptophan residue: the mechanosensitive channel MscL. *Biochemistry* **42**, 14306–14317 (2003).
79. Mitra, K., Ubarretxena-Belandia, I., Taguchi, T., Warren, G. & Engelman, D. M. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl Acad. Sci. USA* **101**, 4083–4088 (2004).
80. Abramovitch, D. A., Marsh, D. & Powell, G. L. Activation of beef-heart cytochrome c oxidase by cardiolipin and analogues of cardiolipin. *Biochim. Biophys. Acta* **1020**, 34–42 (1990).
81. Valiyaveetil, F. I., Zhou, Y. & MacKinnon, R. Lipids in the structure, folding, and function of

- the KcsA K⁺ channel. *Biochemistry* **41**, 10771–10777 (2002).
82. Schonbrun, J., Wedemeyer, W. J. & Baker, D. Protein structure prediction in 2002. *Curr. Opin. Struct. Biol.* **12**, 348–354 (2002).
 83. Briggs, J. A., Torres, J. & Arkin, I. T. A new method to model membrane protein structure based on silent amino acid substitutions. *Proteins* **44**, 370–375 (2001).
 84. Kim, S., Chamberlain, A. K. & Bowie, J. U. Membrane channel structure of *Helicobacter pylori* vacuolating toxin: role of multiple GXXXG motifs in cylindrical channels. *Proc. Natl Acad. Sci. USA* **101**, 5988–5991 (2004).
 85. Sale, K., Faulon, J. L., Gray, G. A., Schoeniger, J. S. & Young, M. M. Optimal bundling of transmembrane helices using sparse distance constraints. *Protein Sci.* **13**, 2613–2627 (2004).
 86. Adams, P., Engelman, D. & Brünger, A. Improved prediction for the structure of the dimeric transmembrane domain of glycophorin A obtained through global searching. *Proteins Struct. Func. Genet.* **26**, 257–261 (1996).
 87. Fleishman, S. J., Unger, V. M., Yeager, M. & Ben-Tal, N. A C α model for the transmembrane α helices of gap junction intercellular channels. *Mol. Cell* **15**, 879–888 (2004).
 88. White, S. H. & Wimley, W. C. Hydrophobic interactions of peptides with membrane interfaces. *Biochim. Biophys. Acta* **1376**, 339–352. (1998).
 89. MacKenzie, K., Prestegard, J. & Engelman, D. A transmembrane helix dimer: structure and implications. *Science* **276**, 131–133 (1997).
- Acknowledgements** I would like to thank members of my lab for helpful comments and NIH for support. J.U.B. is a Leukemia and Lymphoma Society Scholar.
- Author Information** Reprints and permissions is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to J.U.B. (bowie@mbi.ucla.edu).

Membrane curvature and mechanisms of dynamic cell membrane remodelling

Harvey T. McMahon¹ & Jennifer L. Gallop¹

Membrane curvature is no longer seen as a passive consequence of cellular activity but an active means to create membrane domains and to organize centres for membrane trafficking. Curvature can be dynamically modulated by changes in lipid composition, the oligomerization of curvature scaffolding proteins and the reversible insertion of protein regions that act like wedges in membranes. There is an interplay between curvature-generating and curvature-sensing proteins during vesicle budding. This is seen during vesicle budding and in the formation of microenvironments. On a larger scale, membrane curvature is a prime player in growth, division and movement.

Cellular membranes change conformation in striking ways during such processes as movement, division, the extension of neuronal arbors and vesicle trafficking. Vesicle budding and fusion occur with flux constantly maintaining the communication between membrane-bound compartments. In other cases, membrane curvatures are stabilized and are more permanent, for example in microvilli or the dendritic tree. In Fig. 1a we highlight the areas of the cell where there are regions of high membrane curvature.

Dynamic membrane remodelling is achieved by the interplay between lipids and proteins, and in this review we discuss the mechanisms that are used by the cell to generate, sense and stabilize local regions of membrane curvature. Areas of high membrane curvature frequently exist for only limited periods of time, and this is achieved primarily by using surrounding proteins to change the morphology. Thus in the formation of highly curved vesicles, the curvature is induced by the effects of membrane-associated proteins, the 'coat proteins'. The curvature is readily reversible when the coats dissociate, leaving the vesicle more fusogenic (as their curvature is not stabilized) and the coat proteins can now be reused in a further round of vesicle formation (giving an efficiency to protein usage).

Recent studies have shown how the highly dynamic changes in membrane curvature that accompany vesicle trafficking are brought about, and we discuss this emerging field. The topology of a budding vesicle has different degrees of positive and negative curvature (Fig. 2). There are key roles for the insertion of amphipathic helices in generating curvature and for BAR domains in sensing and stabilizing curvature. We introduce the ideas of local curvature generation, and how this is transmitted and maintained over a wider area by stabilizing domains and coat proteins.

We go on to address how membrane subdomains with a given curvature may have precise biological properties. They may lead to spatially regulated clustering of downstream interaction partners, or to the colocalization of transiently interacting proteins on the basis of curvature. Curvature modules within proteins are conjugated with other protein motifs and domains, and from these collaborative activities we can suggest some new ideas for how membrane curvature can be generated by multiple mechanisms and integrated into cell biology. But we start by considering the properties of membranes

and how the lipid and protein components can influence bilayer topology.

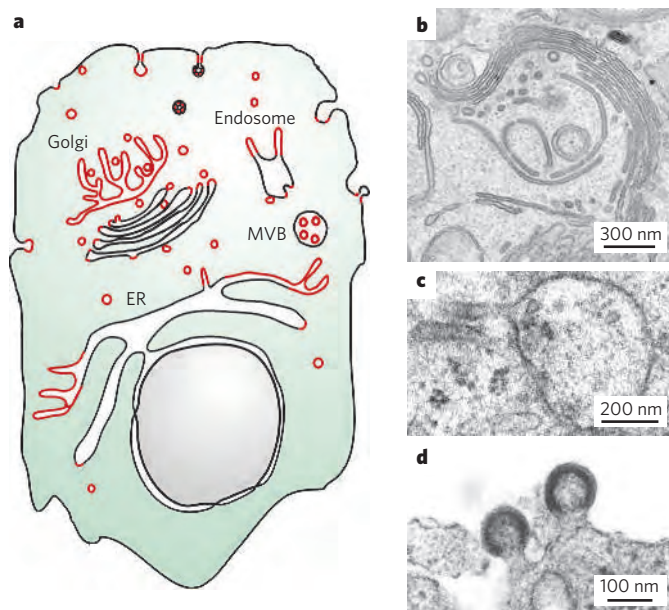


Figure 1 | Local differences in membrane curvature are hallmarks of cellular membranes. Many of the fine details of high local membrane curvature can be seen from the diagram (a) and the sample electron micrographs: b, fenestrations in the Golgi (from C. Hopkins and J. Burden, Imperial College London); c, tubule on endosomes (from P. Luzio and N. Bright, University of Utah); and d, HIV-1 viral budding (from W. Sundquist and U. von Schwedler, University of Utah). All of these can be described as local areas of positive or negative curvature (areas of high positive membrane curvature in a cell highlighted in red). Although it is fascinating to wonder how different membrane morphologies are adapted to the functions of different organelles, we concentrate here on how dynamic changes in morphology are generated. MVB, multi-vesicular body; ER, endoplasmic reticulum.

¹MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

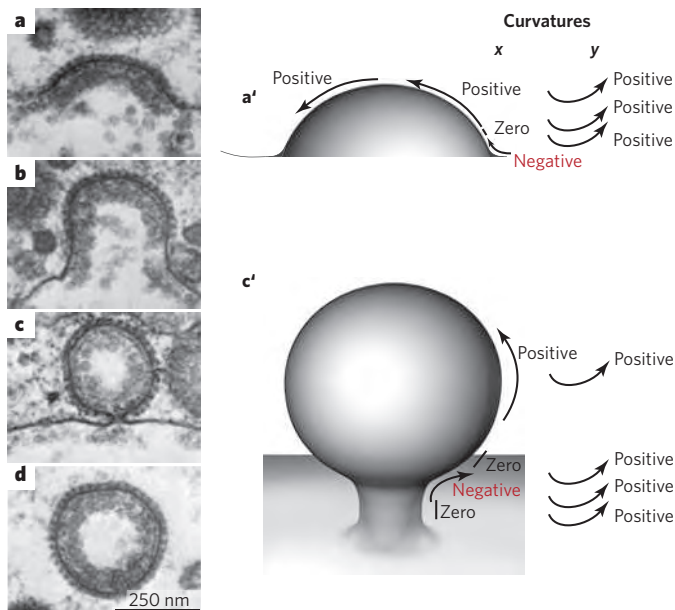


Figure 2 | Clathrin-coated vesicle budding where yolk protein is being incorporated into vesicles in oocytes. (From ref. 88; reproduced with permission from The Company of Biologists.) The different stages (a–d) show progression of invagination and the corresponding membrane curvatures (a', c'). Given that the membrane surface is two-dimensional we need to describe curvature in perpendicular directions. A sphere is positive in both directions, and the curvature of a tubule is positive in one direction and zero in a perpendicular direction. The curvature of the initial stage of vesicle budding is positive in both directions (a'). The curvature of a late-stage budding vesicle is more complex (c'). There is bidirectional positive curvature around the body of the vesicle, negative plus positive curvature (in perpendicular directions) at the neck and interface with the parent membrane, and positive plus zero curvature at the neck of a deeply invaginated vesicle. These types of curvature are constantly being formed and dissolved by the interplay between lipids and proteins.

The lipid component of membranes

The bilayer is a permeability barrier that separates the cell from its exterior and organelles from the cytoplasm. This allows a complex range of reactions both within these compartments and on the membrane surfaces. To communicate between the compartments, vesicles and tubules bud from donor compartments and fuse with others^{1–4}. We may well wonder how these intermediates are formed, as such extreme deformation is unlikely to form spontaneously. The lipids in cell membranes are in a disordered liquid state⁵, meaning that they are free to diffuse and mix in the plane of the bilayer leaflets, although the process may be more complicated than simple brownian diffusion⁶. Lipid mixtures *in vitro* do not readily reconstitute the local variations in curvature of organelle membranes. However, high mole fractions of some lipids are capable of deforming liposomes into tubules⁷, and curvature-gymnastics are seen in giant liposomes of relatively simple lipid compositions, where different lipids segregate according to their chemical properties and lead to the formation of buds and domains on the liposome^{8,9}. Such behaviours seem very 'cell-like' and clearly the lipid component of the membrane is capable of achieving distinct topologies, although the scale of these deformations is much larger than those discussed here. Moreover, the much more complex lipid mixtures present in a biological membrane, the significant protein component and the control that is needed over membrane dynamics mean that proteins have a crucial function in generation of cell-membrane morphology.

Membrane topology

A large portion of this review concerns vesicle trafficking, and so we describe the membrane curvatures that form a budding vesicle. We use

'positive' to indicate regions of membrane that curve inwards towards the cytoplasm. By this definition, the early stages of vesicle budding (shallow pits) have positive curvature and viruses budding out of the cell have negative curvature (see Figs 1 and 2). We will first consider the early stages of budding. The curvature of the dome (Fig. 2a,a') can be described as having positive curvature in two directions. This matures into a deeply invaginated vesicle, which is ready to bud off (Fig. 2c,c'). At this stage there are more components of curvature present than there are in a simple dome. At the transition between the dome of the vesicle and the neck there is both positive and negative curvature, in perpendicular directions. There is positive curvature because the neck is still round (a transverse section would give a circle), but there is also negative curvature because there is formation of a concave surface (a longitudinal section of the budding vesicle shows an omega shape). The neck itself, present to a greater or lesser degree, is shaped like a cylinder; there is still positive curvature in one direction (it is still round) but zero curvature in the other (the sides approximate a straight line).

Five ways to bend a membrane

There are several mechanisms that could generate positive or negative curvature. The following five divisions (see Fig. 3) are used for simplicity, and we do not expect these processes to work in isolation.

Changes in lipid composition

At the very least, lipids have a permissive role in membrane curvature. The chemical properties of different lipid acyl chains or headgroups can favour different membrane curvatures: for example, lysophosphatidic acid (LPA) and phosphatidic acid (PA), which are interconverted by lysophosphatidic acid acyl transferase and phospholipase A₂ activity respectively^{10–12}, favour opposite curvatures. In addition, flippases (which transfer lipids from one leaflet to the other) give rise to membrane asymmetry^{13,14}, and enzymes that change lipid headgroup size will influence the area occupied by the lipids¹⁵ and thus affect membrane curvature. Some of these changes may well be localized by limited diffusion barriers (for example the presence of transmembrane proteins or the knitting together of proteins by cytoskeletal or scaffolding attachments), and thus they may assist or antagonize changes in topology.

Lipid headgroups are the attachment sites for peripheral membrane proteins and therefore aid the recruitment of proteins necessary to generate curvature. Phosphoinositides (PtdIns) are particularly important as their headgroups are easily modified (see the review by Behnia and Munro in this issue, p. 597). For example, the presence of PtdIns(4,5)P₂ in the plasma membrane is essential for the budding of clathrin-coated vesicles, largely because the budding machinery binds to PtdIns(4,5)P₂ (refs 16–20). Similarly, in the invagination of vesicles into late endosomes there is a requirement for PtdIns-3-OH kinase²¹, with which Hrs and other FYVE domain proteins interact²², and for PtdIns(3)P-5-OH kinase²³.

As well as assisting or antagonizing curvature, lipids may also respond to curvature by concentrating in domains of curvature that they prefer. This is seen in the tubule-pulling experiment of the Goud laboratory²⁴ where lipids segregate into the tubules.

Inherent in the small size (and high curvature) of transport intermediates is an imbalance in the number of lipids in the inner and outer leaflets of the bilayer. In a liposome with an outer diameter of 50 nm with a membrane thickness of 5 nm, there is 56% more lipid in the outer leaflet than in the inner. So when a vesicle fuses, the imbalance in the outer and inner leaflet lipids has to be accommodated or there has to be a compensatory change. For example, the generation of negative curvature at the neck of the vesicle will, at least in part, relax the positive curvature of the dome.

When making dynamic or reversible changes in membrane curvature (as in making a transport vesicle or tubule) it may well be advantageous to avoid giving long-term stability to the high membrane curvatures as these trafficking intermediates will need to fuse with

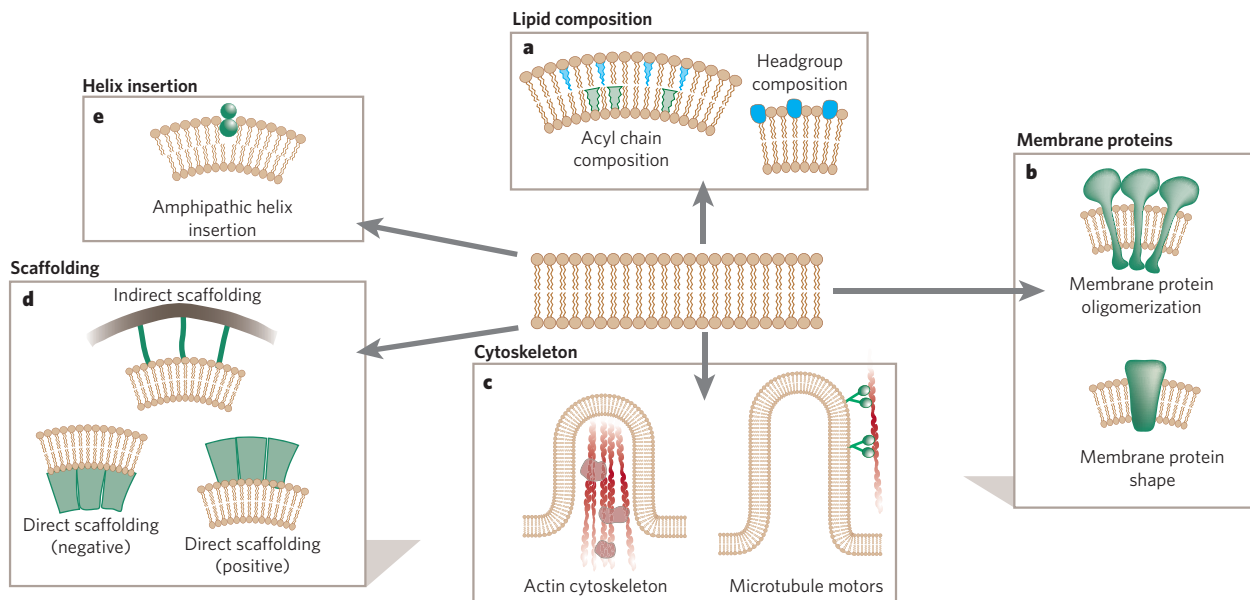


Figure 3 | Mechanisms of membrane deformation. The phospholipid bilayer can be deformed causing positive or negative membrane curvature. There are five main categories: **a**, changes in lipid composition; **b**, influence of integral membrane proteins that have intrinsic curvature or have curvature on oligomerization; **c**, changes in cytoskeletal polymerization and pulling of tubules by motor proteins; **d**, direct and indirect scaffolding of the bilayer; **e**, active amphipathic helix insertion into one leaflet of the bilayer.

recipient membranes, and this process may be aided by the instability and tension inherent in the high curvature. In such cases, peripheral protein association would be the primary driver of curvature, although timed headgroup turnover (as in the hydrolysis of PtdIns(4,5)P₂ by synaptojanin²⁵) could also participate.

Influence of integral membrane proteins

Transmembrane proteins with a conical shape will naturally prefer curvatures that mould around their shapes. This shape is seen for the transmembrane domain of the nicotinic acetylcholine receptor, which has been observed at the tops of membrane folds at the neuromuscular junction^{26,27} and is also seen in the structure of the voltage-dependent K⁺-channel²⁸. Acetylcholine receptors and many other transmembrane receptors and channels can be clustered by attachment proteins^{29,30}, leading to a greater effect on local curvature. If the membrane-spanning domain itself is not funnel shaped then curvature could theoretically still be caused by the overall conformation of clustered proteins or a conformational change, perhaps in response to ligand binding. Given that the structures of so few transmembrane proteins are known, the contribution of intrinsic shape to membrane curvature localization is a virgin field. It would be interesting if receptors destined for endocytosis were to partition and concentrate into regions of high positive curvature (leading to the exclusion of receptors not to be trafficked) or even aid the induction of curvature by lowering the energetic requirements. Indeed, progression of coated pits into vesicles occurs in tandem with cargo loading³¹. The role of curvature in defining membrane domains and in ion channel activity and receptor activation remains largely unexplored and has potential for new insights^{32–34}.

Cytoskeletal proteins and microtubule motor activity

Cytoskeletal assembly and disassembly is intimately linked with membrane-shape changes of the plasma membrane and of organelles^{35,36}. Branching, bundling and treadmilling of actin filaments are involved in the generation and remodelling of many areas of high membrane curvature, including filopodia, pseudopodia, phagocytic cups and axonal growth cones. The ability of the cytoskeleton to influence membrane-shape changes is affected by membrane tension³⁷, and decreases in tension can help the generation of local curvature (for example, membrane trafficking events^{13,38–40}). The cytoskeleton has a large role

in maintaining membrane tension, and conversely actin rearrangements are responsive to changes in tension⁴¹. Therefore, we would envisage constant interplay between the responsive and propulsive power of the cytoskeleton and all the other factors that influence membrane tension and curvature, including trafficking and cell–cell adhesion.

Bursts of actin polymerization have been implicated in many endocytic invagination events^{42–46}. Because actin polymerization has a force-generating role during motility and phagocytosis, it is tempting to assume that the reason for it here is the same, in aiding fission⁴³, but this is not yet clear.

In vivo imaging of cells shows that many tubules and vesicles move along microtubule tracks⁴⁷. *In vitro* it can be demonstrated that kinesin motors attached to Golgi membranes can pull out tubules, and this can be achieved from liposomes with a modest number of motors²⁴. Thus it is very likely that motors are at least partly responsible for fenestrated or tubulated organelle morphology (for example, the ER, Golgi and endosome) and the generation of some transport intermediates^{48,49}. Given the evidence in favour of microtubules in vesicle generation, it is also possible that actin has a similar role with transport of vesicles mediated by myosins⁵⁰.

It is not surprising that cytoskeletal changes influence membrane remodelling in cell motility⁵¹ and in tubule and vesicle carrier formation^{52,53}, but the cytoskeleton also has another function in directing the location of fusing and endocytosing vesicles and in localizing receptors and signalling complexes⁵⁴. Also, many BAR-domain proteins have links by way of signalling proteins to the actin- and microtubule-polymerization machinery (for example tuba, β -centaurins and nadrins⁵⁵; see also <http://www.endocytosis.org/BARdomains/BARs.html>). Much future interest will certainly lie in this interface between the cytoskeleton and the proteins that sense or drive curvature.

Scaffolding by peripheral membrane proteins

This can take different forms. Here we consider several families of proteins that deform a membrane by bracing it like a scaffold.

Proteins of the dynamin family bind to inositol lipids and form helical oligomers, constraining the membrane topology into a tubular shape^{56,57}. They have an important role in the constriction of organelles during their division, in forming the necks of invaginating vesicles and

promoting their scission from the parent membrane. In plants they generate tubules during cell-wall formation⁵⁸. This family of proteins uses GTP hydrolysis to effect membrane fission^{57,59}. An analogy for the role of these proteins is an exoskeleton, supporting and sculpting the membrane from the outside. This can also be achieved by an endoskeleton, as in viral budding⁶⁰.

Coat proteins such as clathrin, COPI and COPII can also be considered as exoskeletons that influence membrane bending by polymerizing into curved structures, but these coat proteins do not have direct membrane associations and so are likely to act in conjunction with other proteins (see below)^{61,62}. Caveoli are flask-shaped membrane invaginations where caveolin oligomerizes to form the coat⁶³. Unlike COP and clathrin-coated vesicles, caveolin is membrane-associated and this could aid in membrane bending by insertion.

BAR domains are modules that sense membrane curvature (see Box 1). This ability to bind preferentially to curved membranes can be deduced from the concave shape of the membrane-binding region. The sensing is shown by its tighter binding to liposomes whose curvature is closer to the intrinsic curvature of the BAR⁵⁵. The energetics of BAR-domain binding to membranes for amphiphysin also leads to the conclusion that the binding energy is close to that required to bend the membrane⁶⁴.

BAR domains are formed by dimerization, which is probably enhanced by membrane binding, and therefore the other constituent domains of the protein are presented as pairs. This could, for example, lead to the co-recruitment of two binding partners or a change in selection of a monomeric for dimeric/multimeric partners and thus generate a unique downstream signal based on the initial curvature-sensitive binding event.

BAR domains are also frequently found in combination with N-terminal amphipathic helices (Box 2). They are then called N-BAR domains (see below). This is an interesting combination and can be seen in amphiphysin, endophilin, BRAP and nadrin. All these N-BAR domains lead to membrane tubulation *in vitro*^{55,65–68}. In *Drosophila* the N-BAR protein amphiphysin is involved in T-tubule formation in flight muscles and in its absence the T-tubule network is disrupted, preventing flight. In the synapse, amphiphysin is proposed to form or stabilize a very different tubule structure, that of the neck of clathrin-coated vesicles. The degree of positive curvature of the neck is close to that of the BAR, and thus this protein is suited for the recruitment of its binding partner, dynamin, to its correct location^{65,69}. Dynamin may also aid in neck formation as it polymerizes into tubules of the same diameter (see exoskeleton discussion above). BAR domains and homologous domains are found in many trafficking proteins and their role in curvature sensing and stabilization will need much more study. We have recently shown that the BAR-domain protein sorting nexin-1 is involved in tubule extension from endosomes⁷⁰. This protein seems to coat the tubule extensions that are involved in trafficking mannose-6-phosphate receptors to the *trans*-Golgi network (TGN).

Active helix insertion into membranes

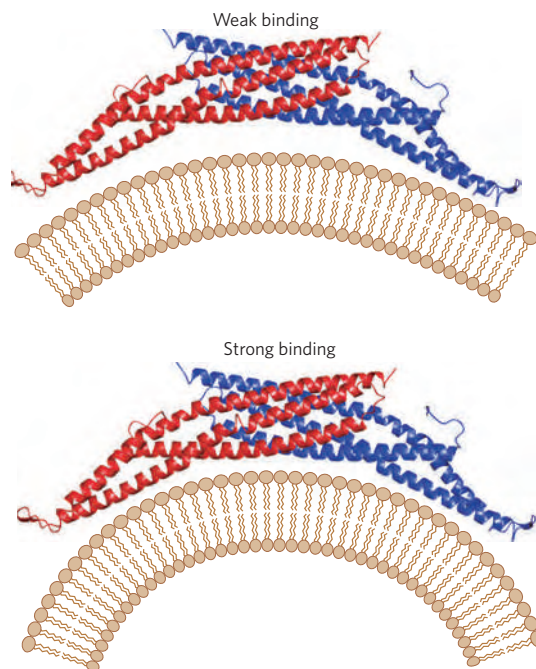
Amphipathic helices inserted into the bilayer result in increased positive membrane curvature (Box 2). In the case of epsin this helix folds and inserts on PtdIns(4,5)P₂ binding. Epsin in turn binds to clathrin and promotes its polymerization into a cage-like structure, and this stabilizes the change in local curvature. Amphiphysin, endophilins, BRAPs and nadrins all have BAR domains with an amphipathic helix at the N terminus (N-BAR domains). These should work in a similar manner⁶⁸ to cause local membrane curvature and in these cases we would predict curvature stabilization by the banana-shaped BAR domain instead of by clathrin or another coat protein. Arf and Arl proteins also have N-terminal amphipathic helices that are extended in response to GTP binding, and Arfs are involved in COP1 vesicle budding and in recruitment of GGA and AP1 complexes to membranes^{71–73}. By analogy with epsins, these Arfs and Arls are predicted to function in curvature generation alongside stabilization by coat proteins. Sar1 (another small GTPase with an N-terminal amphipathic

Box 1 | BAR domains and stabilization of membrane curvature

BAR domains are banana-shaped lipid-binding domains found in a wide variety of proteins, which bind to membranes through their concave surface⁵⁶ (see also <http://www.endocytosis.org/BARdomains/BARs.html>). They are dimers, and given that the dimer interface and the membrane-binding region overlap, membrane binding may stabilize the dimer formation⁸⁹. If dimerization is more effective on membrane binding than in the cytosol then multimeric effectors will be better recruited to a membrane-bound protein. For example, dynamin (which is a dimer) binding to amphiphysin will clearly be of much higher avidity when amphiphysin is a dimer.

The BAR interaction with membranes is largely electrostatic and binds to negatively charged membranes. A high concentration of lysine and arginine residues between helices 2 and 3 in some BARs help to give some PtdIns(4,5)P₂ preference over PtdSer (see also the review by McLaughlin and Murray in this issue, p. 605). Other BAR proteins contain specific membrane-targeting PH or PX domains to locate them to specific compartments⁵⁶. BAR domains bind more readily to highly curved liposomes (see Box 1 Fig. 1)⁵⁶. Thus the domain on its own is a sensor of high positive curvature. We should also consider that given a high concentration of a curvature sensor it is clearly possible that a sensor will become an inducer.

An additional feature of some BAR domains is the presence of an N-terminal amphipathic helix (an N-BAR domain). As discussed in Box 2, this amphipathic helix will lead to membrane bending. Thus it is interesting to find these two curvature modules side by side in many proteins.



Box 1 Figure 1 | The amphiphysin BAR domain in association with low-curvature and high-curvature membranes. The BAR domain binds better to the more highly curved membranes because there is more opportunity for electrostatic interactions across the complete membrane-binding surface of the BAR.

helix) is likely to function in a similar manner for COPII-coated vesicle budding. The COPII coat structure has already been shown to have a surface that will follow the curved membrane and thus stabilize the curvature⁷⁴.

Coupling curvature to function

The examples below illustrate the involvement of lipids and proteins in the formation of positive and negative membrane curvature. We concentrate on the making of transport vesicles where membrane curvature is mediated by the collaboration of different mechanisms at different stages of budding events. The lessons can be extended to the

Amphipathic helices are stretches of α -helix, one side of which is polar (charged) and the other hydrophobic. These helices are frequently unstructured until they insert into membranes, when the helices are predicted to sit flat on the membrane surface with the hydrophobic residues dipping into the hydrophobic phase of the membrane^{17,55}. The result will be a displacement of lipid headgroups and a reorientation of acyl chains, giving an orientation more favourable to higher curvature. The fact that this mechanism of curvature can be reproduced on a lipid monolayer¹⁸ shows that this is not just a generation of bilayer asymmetry, nor simply headgroup displacement, but that it is primarily the reordering of the lipids in an individual leaflet with a tighter bend. We can model how the helix looks by taking an ideal α -helix and modelling the sequence of interest. In the figure we illustrate this for the initial residues of *Drosophila* amphiphysin, which *in vivo* is involved in stabilization of T-tubule formation in muscles⁶⁷. This model is an oversimplification as there is sometimes a kink in the helix and the nature of the polar face may give different properties to these helices (see work on synuclein and on ArfGAP1 and synthetic peptides^{78,89,90}). We have previously shown that the N-terminal residues of amphiphysin adopt a helical conformation on membrane binding and it is clear that from residue 9 forward there is a strong hydrophobic face and a polar face. Another way to visualize the amphipathic nature of a stretch of amino acids is to use an axial projection of the helix (a helical wheel, see <http://www.site.uottawa.ca/~turtotte/resources/HelixWheel/>).

The most important feature of the amphipathic helix for this review is its effect on membrane curvature. Given the asymmetric insertion (see figure) it acts like a wedge inserted into one leaflet of the membrane. All the amphipathic helices we have studied effect membrane curvature given a high local concentration. Thus it makes sense that epsins bind and promote clathrin polymerization, concentrating the curvature into a local membrane area.

Polar/charged face

Hydrophobic face

Initial stage of amphipathic helix folding and insertion

Membrane curvature response

d-Amph 3-20

We have argued that amphipathic helices will promote an increase in membrane curvature when folded and inserted between the lipid headgroups, but it is also entirely possible that some amphipathic helices will insert only in response to high curvature, and thus even a humble helix may act like a curvature sensor. This is likely to be the case for ArfGAP1, which promotes Arf1 GTP hydrolysis during COPI vesicle budding, thus coupling vesicle budding to the initial stages of the uncoating reaction⁷⁸.

©2005 Nature Publishing Group

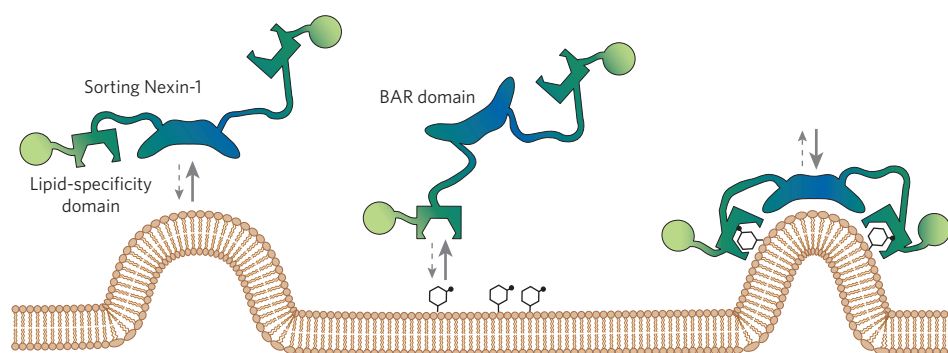


Figure 4 | Coincidence detection. In sorting nexin-1 there is a dimerized BAR domain (blue), which recognizes membrane curvature, and an additional lipid specificity domain (dark green, PX domain). Binding of both domains — that is, coincidence detection of both membrane composition and curvature — is required for recruitment of the protein and for stabilization of membrane curvature. Hexagons represent phosphatidylinositol phosphate headgroups to which PX domains bind.

mal compartment, presumably because of an inability to invaginate the limiting membrane to form the MVB. These proteins can be sorted into complexes and associated proteins that together dynamically interact with endosomal membranes. They help sort cargo and may well have an active role in the exagination process. Alix/Bro1, a class E Vps mutant involved in MVB formation⁸³, binds to LBPA-containing liposomes and regulates the formation of internal vesicles⁸³. The structure of a Bro1 domain has an interesting boomerang shape⁸⁴ (somewhat like the BAR domain) and could potentially function in negative curvature generation, but there is as yet little evidence for this. RNA interference of LIP5 and CHMP4 inhibit MVB formation and also the budding of HIV-1 viral particles⁸⁵. Given the number of coiled-coil proteins involved in MVB formation, it would be interesting if there were an inverse BAR domain, an 'I-BAR'.

Future perspectives

Coincidence detection and membrane microenvironments

A cell has many curved membranes (Fig. 1) and so additional mechanisms of selection must be used if the cell is to sense and respond selectively to membrane curvature. The coincident detection of a number of inputs is a common theme in biology that minimizes noise and gives highly selective responses. BAR-domain proteins give us an example of coincidence detection between lipid curvature and composition. We have already noted the presence of PH domain or PX domains alongside BAR domains in the same proteins⁵⁵. Point mutants show that the PH domains, the PX domains or the BAR domains alone are insufficient for membrane targeting but that these domains work together (Fig. 4)^{55,70}. This gives rise to a precise localization of the protein in question. Given the range of domains found in proteins with BAR modules it is likely that this coincidence detection will work for many different levels. Again we can use an example from proteins containing BAR domains. Some of those so far identified have GAP and GEF activities. By analogy to the above examples with PX or PH domains, it can be predicted that these may be curvature sensitive. What seemed remarkable about the PH-BAR and PX-BAR examples is that the proteins do not target visibly in the absence of either domain. The BAR domain lipid-binding mutant was not expected to disrupt dimerization of the BAR, but surprisingly the supposedly dimeric protein (thus two PH domains) does not localize to membranes. This may hint at the importance of membrane binding for stable dimer formation, or it may simply mean that the BAR and the two PH domains are all required for membrane binding.

The selective binding of proteins depending on curvature and the partitioning of lipids favouring that curvature into these regions gives the exciting possibility of a local microenvironment on the membrane. It could, for example, favour the segregation of transmembrane proteins for incorporation into vesicles or tubules or the preferential localization of ion channels in protrusions. It would be interesting to know whether the transport tubules extending from endosomes could concentrate cargo by curvature preference. Similarly, the localization of GAP or GEF activities according to curvature could lead to tight regulation of small G-protein GTP/GDP status and therefore (for instance) selective actin polymerization or signalling pathway activa-

tion at these domains. This is like an ecological niche, where curvature defines a protein-lipid microenvironment in which specific interactions are more likely to occur. This could be a dynamic environment where the domain is transient and only forms in response to a range of coincident stimuli.

The interplay between lipids and proteins is key to how cells control membrane shape. This ability of proteins to alter membrane curvature directly is an emerging field of study and the above discussions readily illustrate the importance of multiple mechanisms to obtain effective membrane curvature changes. As in clathrin-coated vesicle formation there is a network of interactions⁸⁶ and interlinking pathways that must be considered before we will have understood how cells generate, control and respond to curvature domains and dynamics. ■

- Rothman, J. E. & Orci, L. Budding vesicles in living cells. *Scient. Am.* **274**, 70–75 (1996).
- Sciaky, N. *et al.* Golgi tubule traffic and the effects of brefeldin A visualized in living cells. *J. Cell Biol.* **139**, 1137–1155 (1997).
- Heuser, J. Three-dimensional visualization of coated vesicle formation in fibroblasts. *J. Cell Biol.* **84**, 560–583 (1980).
- Bright, N. A., Gratian, M. J. & Luzio, J. P. Endocytic delivery to lysosomes mediated by concurrent fusion and kissing events in living cells. *Curr. Biol.* **15**, 360–365 (2005).
- Singer, S. J. & Nicolson, G. L. The fluid mosaic model of the structure of cell membranes. *Science* **175**, 720–731 (1972).
- Kusumi, A. *et al.* Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: high-speed single-molecule tracking of membrane molecules. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 351–378 (2005).
- Stowell, M. H., Marks, B., Wigge, P. & McMahon, H. T. Nucleotide-dependent conformational changes in dynamin: evidence for a mechanochemical molecular spring. *Nature Cell Biol.* **1**, 27–32 (1999).
- Baumgart, T., Hess, S. T. & Webb, W. W. Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Nature* **425**, 821–824 (2003).
- Bacia, K., Schwill, P. & Kurzchalia, T. Sterol structure determines the separation of phases and the curvature of the liquid-ordered phase in model membranes. *Proc. Natl Acad. Sci. USA* **102**, 3272–3277 (2005).
- Kooijman, E. E. *et al.* Spontaneous curvature of phosphatidic acid and lysophosphatidic acid. *Biochemistry* **44**, 2097–2102 (2005).
- Brown, W. J., Chambers, K. & Doody, A. Phospholipase A2 (PLA₂) enzymes in membrane trafficking: mediators of membrane shape and function. *Traffic* **4**, 214–221 (2003).
- Shemesh, T., Luini, A., Malhotra, V., Burger, K. N. & Kozlov, M. M. Prefission constriction of Golgi tubular carriers driven by local lipid metabolism: a theoretical model. *Biophys. J.* **85**, 3813–3827 (2003).
- Farage, E., Ojcius, D. M., Subtil, A. & Dautry-Varsat, A. Enhancement of endocytosis due to aminophospholipid transport across the plasma membrane of living cells. *Am. J. Physiol.* **276**, C725–C733 (1999).
- Hua, Z. & Graham, T. R. Requirement for neo1p in retrograde transport from the Golgi complex to the endoplasmic reticulum. *Mol. Biol. Cell* **14**, 4971–4983 (2003).
- Hammond, K., Reboiras, M. D., Lyle, I. G. & Jones, M. N. Characterisation of phosphatidylcholine/phosphatidylinositol sonicated vesicles. Effects of phospholipid composition on vesicle size. *Biochim. Biophys. Acta* **774**, 19–25 (1984).
- Ford, M. G. *et al.* Simultaneous binding of PtdIns(4,5)P₂ and clathrin by AP180 in the nucleation of clathrin lattices on membranes. *Science* **291**, 1051–1055 (2001).
- Ford, M. G. *et al.* Curvature of clathrin-coated pits driven by epsin. *Nature* **419**, 361–366 (2002).
- Kinuta, M. *et al.* Phosphatidylinositol 4,5-bisphosphate stimulates vesicle formation from liposomes by brain cytosol. *Proc. Natl Acad. Sci. USA* **99**, 2842–2847 (2002).
- Wenk, M. R. & De Camilli, P. Protein-lipid interactions and phosphoinositide metabolism in membrane traffic: insights from vesicle recycling in nerve terminals. *Proc. Natl Acad. Sci. USA* **101**, 8262–8269 (2004).
- Honing, S. *et al.* Phosphatidylinositol-(4,5)-bisphosphate regulates sorting signal recognition by the clathrin-associated adaptor complex AP2. *Mol. Cell* **18**, 519–531 (2005).
- Fernandez-Borja, M. *et al.* Multivesicular body morphogenesis requires phosphatidylinositol 3-kinase activity. *Curr. Biol.* **9**, 55–58 (1999).
- Gruenberg, J. & Stenmark, H. The biogenesis of multivesicular endosomes. *Nature Rev. Mol. Cell Biol.* **5**, 317–323 (2004).

23. Odorizzi, G., Babst, M. & Emr, S. D. Fab1p PtdIns(3)P 5-kinase function essential for protein sorting in the multivesicular body. *Cell* **95**, 847–858 (1998).
24. Roux, A. *et al.* Role of curvature and phase transition in lipid sorting and fission of membrane tubules. *EMBO J.* **24**, 1537–1545 (2005).
25. Cremona, O. *et al.* Essential role of phosphoinositide metabolism in synaptic vesicle recycling. *Cell* **99**, 179–188 (1999).
26. Fertuck, H. C. & Salpeter, M. M. Localization of acetylcholine receptor by ¹²⁵I-labeled alpha-bungarotoxin binding at mouse motor endplates. *Proc. Natl Acad. Sci. USA* **71**, 1376–1378 (1974).
27. Unwin, N. Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *J. Mol. Biol.* **346**, 967–989 (2005).
28. Mackinnon, R. Structural biology. Voltage sensor meets lipid membrane. *Science* **306**, 1304–1305 (2004).
29. Boudin, H. *et al.* Presynaptic clustering of mGluR7a requires the PICK1 PDZ domain binding site. *Neuron* **28**, 485–497 (2000).
30. Eckler, S. A., Kuehn, R. & Gautam, M. Deletion of N-terminal rapsyn domains disrupts clustering and has dominant negative effects on clustering of full-length rapsyn. *Neuroscience* **131**, 661–670 (2005).
31. Kirchhausen, T., Boll, W., van Oijen, A. & Ehrlich, M. Single-molecule live-cell imaging of clathrin-based endocytosis. *Biochem. Soc. Symp.* **72**, 71–76 (2005).
32. Petrou, S. *et al.* Direct effects of fatty acids and other charged lipids on ion channel activity in smooth muscle cells. *Prostaglandins Leukot. Essent. Fatty Acids* **52**, 173–178 (1995).
33. Casado, M. & Ascher, P. Opposite modulation of NMDA receptors by lysophospholipids and arachidonic acid: common features with mechanosensitivity. *J. Physiol.* **513**, 317–330 (1998).
34. Fuster, D., Moe, O. W. & Hilgemann, D. W. Lipid- and mechanosensitivities of sodium/hydrogen exchangers analyzed by electrical methods. *Proc. Natl Acad. Sci. USA* **101**, 10482–10487 (2004).
35. Ledesma, M. D. & Dotti, C. G. Membrane and cytoskeleton dynamics during axonal elongation and stabilization. *Int. Rev. Cytol.* **227**, 183–219 (2003).
36. Sheetz, M. P. Cell control by membrane-cytoskeleton adhesion. *Nature Rev. Mol. Cell Biol.* **2**, 392–396 (2001).
37. Raucher, D. & Sheetz, M. P. Cell spreading and lamellipodial extension rate is regulated by membrane tension. *J. Cell Biol.* **148**, 127–136 (2000).
38. Dai, J., Ting-Beall, H. P. & Sheetz, M. P. The secretion-coupled endocytosis correlates with membrane tension changes in RBL 2H3 cells. *J. Gen. Physiol.* **110**, 1–10 (1997).
39. Heidelberger, R., Zhou, Z. Y. & Matthews, G. Multiple components of membrane retrieval in synaptic terminals revealed by changes in hydrostatic pressure. *J. Neurophysiol.* **88**, 2509–2517 (2002).
40. Raucher, D. & Sheetz, M. P. Membrane expansion increases endocytosis rate during mitosis. *J. Cell Biol.* **144**, 497–506 (1999).
41. Bettache, N. *et al.* Mechanical constraint imposed on plasma membrane through transverse phospholipid imbalance induces reversible actin polymerization via phosphoinositide 3-kinase activation. *J. Cell Sci.* **116**, 2277–2284 (2003).
42. Merrifield, C. J., Perraies, D. & Zenisek, D. Coupling between clathrin-coated-pit invagination, cortactin recruitment, and membrane scission observed in live cells. *Cell* **121**, 593–606 (2005).
43. Yasar, D., Waterman-Storer, C. M. & Schmid, S. L. A dynamic actin cytoskeleton functions at multiple stages of clathrin-mediated endocytosis. *Mol. Biol. Cell* **16**, 964–975 (2005).
44. Shupliakov, O. *et al.* Impaired recycling of synaptic vesicles after acute perturbation of the presynaptic actin cytoskeleton. *Proc. Natl Acad. Sci. USA* **99**, 14476–14481 (2002).
45. Engqvist-Goldstein, A. E. *et al.* RNAi-mediated Hip1R silencing results in stable association between the endocytic machinery and the actin assembly machinery. *Mol. Biol. Cell* **15**, 1666–1679 (2004).
46. Qualmann, B. & Kelly, R. B. Synaptic isoforms participate in receptor-mediated endocytosis and actin organization. *J. Cell Biol.* **148**, 1047–1062 (2000).
47. Rodriguez-Boulant, E., Kreitzer, G. & Musch, A. Organization of vesicular trafficking in epithelia. *Nature Rev. Mol. Cell Biol.* **6**, 233–247 (2005).
48. Vale, R. D. & Hotani, H. Formation of membrane networks in vitro by kinesin-driven microtubule movement. *J. Cell Biol.* **107**, 2233–2241 (1988).
49. Dabora, S. L. & Sheetz, M. P. The microtubule-dependent formation of a tubulovesicular network with characteristics of the ER from cultured cell extracts. *Cell* **54**, 27–35 (1988).
50. Buss, F., Luzzio, J. P. & Kendrick-Jones, J. Myosin VI, an actin motor for membrane traffic and cell migration. *Traffic* **3**, 851–858 (2002).
51. Bretscher, M. S. Getting membrane flow and the cytoskeleton to cooperate in moving cells. *Cell* **87**, 601–606 (1996).
52. Allan, V. & Vale, R. Movement of membrane tubules along microtubules in vitro: evidence for specialised sites of motor attachment. *J. Cell Sci.* **107**, 1885–1897 (1994).
53. Merrifield, C. J. Seeing is believing: imaging actin dynamics at single sites of endocytosis. *Trends Cell Biol.* **14**, 352–358 (2004).
54. Zakharenko, S. & Popov, S. Dynamics of axonal microtubules regulate the topology of new membrane insertion into the growing neurites. *J. Cell Biol.* **143**, 1077–1086 (1998).
55. Peter, B. J. *et al.* BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* **303**, 495–499 (2004).
56. Hinshaw, J. E. & Schmid, S. L. Dynamin self-assembles into rings suggesting a mechanism for coated vesicle budding. *Nature* **374**, 190–192 (1995).
57. Marks, B. *et al.* GTPase activity of dynamin and resulting conformation change are essential for endocytosis. *Nature* **410**, 231–235 (2001).
58. Praefcke, G. J. & McMahon, H. T. The dynamin superfamily: universal membrane tubulation and fission molecules? *Nature Rev. Mol. Cell Biol.* **5**, 133–147 (2004).
59. Sweitzer, S. M. & Hinshaw, J. E. Dynamin undergoes a GTP-dependent conformational change causing vesiculation. *Cell* **93**, 1021–1029 (1998).
60. von Schwedler, U. K. *et al.* The protein network of HIV budding. *Cell* **114**, 701–713 (2003).
61. Antony, B., Gounon, P., Schekman, R. & Orci, L. Self-assembly of minimal COPII cages. *EMBO Rep.* **4**, 419–424 (2003).
62. Nossal, R. Energetics of clathrin basket assembly. *Traffic* **2**, 138–147 (2001).
63. Razani, B. & Lisanti, M. P. Caveolins and caveolae: molecular and functional relationships. *Exp. Cell Res.* **271**, 36–44 (2001).
64. Zimmerberg, J. & McLaughlin, S. Membrane curvature: how BAR domains bend bilayers. *Curr. Biol.* **14**, R250–R252 (2004).
65. Takei, K., Slepnev, V. I., Haucke, V. & De Camilli, P. Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis. *Nature Cell Biol.* **1**, 33–39 (1999).
66. Farsad, K. *et al.* Generation of high curvature membranes mediated by direct endophilin bilayer interactions. *J. Cell Biol.* **155**, 193–200 (2001).
67. Razaq, A. *et al.* Amphiphysin is necessary for organization of the excitation-contraction coupling machinery of muscles, but not for synaptic vesicle endocytosis in *Drosophila*. *Genes Dev.* **15**, 2967–2979 (2001).
68. Richnau, N., Fransson, A., Farsad, K. & Aspenstrom, P. RICH-1 has a BIN/Amphiphysin/Rvsp domain responsible for binding to membrane lipids and tubulation of liposomes. *Biochem. Biophys. Res. Commun.* **320**, 1034–1042 (2004).
69. Wigge, P. *et al.* Amphiphysin heterodimers: potential role in clathrin-mediated endocytosis. *Mol. Biol. Cell* **8**, 2003–2015 (1997).
70. Carlton, J. *et al.* Sorting nexin-1 mediates tubular endosome-to-TGN transport through coincidence sensing of high-curvature membranes and 3-phosphoinositides. *Curr. Biol.* **14**, 1791–1800 (2004).
71. Orcl, L., Palmer, D. J., Amherdt, M. & Rothman, J. E. Coated vesicle assembly in the Golgi requires only coatamer and ARF proteins from the cytosol. *Nature* **364**, 732–734 (1993).
72. Seaman, M. N., Sowerby, P. J. & Robinson, M. S. Cytosolic and membrane-associated proteins involved in the recruitment of AP-1 adaptors onto the trans-Golgi network. *J. Biol. Chem.* **271**, 25446–25451 (1996).
73. Puertollano, R., Randazzo, P. A., Presley, J. F., Hartnell, L. M. & Bonifacio, J. S. The GGAs promote ARF-dependent recruitment of clathrin to the TGN. *Cell* **105**, 93–102 (2001).
74. Bi, X., Corpina, R. A. & Goldberg, J. Structure of the Sec23/24-Sar1 pre-budding complex of the COPII vesicle coat. *Nature* **419**, 271–277 (2002).
75. Mashl, R. J. & Bruinsma, R. F. Spontaneous-curvature theory of clathrin-coated membranes. *Biophys. J.* **74**, 2862–2875 (1998).
76. Kozlov, M. M. Fission of biological membranes: interplay between dynamin and lipids. *Traffic* **2**, 51–65 (2001).
77. Bigay, J., Gounon, P., Robineau, S. & Antony, B. Lipid packing sensed by ArfGAP1 couples COPI coat disassembly to membrane bilayer curvature. *Nature* **426**, 563–566 (2003).
78. Bigay, J., Casella, J. F., Drin, G., Mesmin, B. & Antony, B. ArfGAP1 responds to membrane curvature through the folding of a lipid packing sensor motif. *EMBO J.* **24**, 2244–2253 (2005).
79. Antony, B., Beraud-Dufour, S., Chardin, P. & Chabre, M. N-terminal hydrophobic residues of the G-protein ADP-ribosylation factor-1 insert into membrane phospholipids upon GDP to GTP exchange. *Biochemistry* **36**, 4675–4684 (1997).
80. Kobayashi, T. *et al.* A lipid associated with the antiphospholipid syndrome regulates endosome structure and function. *Nature* **392**, 193–197 (1998).
81. Matsu, H. *et al.* Role of LBPA and Alix in multivesicular liposome formation and endosome organization. *Science* **303**, 531–534 (2004).
82. Katzmman, D. J., Babst, M. & Emr, S. D. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell* **106**, 145–155 (2001).
83. Odorizzi, G., Katzmman, D. J., Babst, M., Audhya, A. & Emr, S. D. Bro1 is an endosome-associated protein that functions in the MVB pathway in *Saccharomyces cerevisiae*. *J. Cell Sci.* **116**, 1893–1903 (2003).
84. Kim, J. *et al.* Structural basis for endosomal targeting by the bro1 domain. *Dev. Cell* **8**, 937–947 (2005).
85. Ward, D. M. *et al.* The role of LIP5 and CHMP5 in multivesicular body formation and HIV-1 budding in mammalian cells. *J. Biol. Chem.* **280**, 10548–10555 (2005).
86. Praefcke, G. J. *et al.* Evolving nature of the AP2 alpha-appendage hub during clathrin-coated vesicle endocytosis. *EMBO J.* **23**, 4371–4383 (2004).
87. Perry, M. M. & Gilbert, A. B. Yolk transport in the ovarian follicle of the hen (*Gallus domesticus*): lipoprotein-like particles at the periphery of the oocyte in the rapid growth phase. *J. Cell Sci.* **39**, 257–272 (1979).
88. Gallop, J. L. & McMahon, H. T. BAR domains and membrane curvature: bringing your curves to the BAR. *Biochem. Soc. Symp.* **72**, 223–231 (2005).
89. Jao, C. C., Der-Sarkissian, A., Chen, J. & Langen, R. Structure of membrane-bound alpha-synuclein studied by site-directed spin labeling. *Proc. Natl Acad. Sci. USA* **101**, 8331–8336 (2004).
90. Lee, S. *et al.* De novo-designed peptide transforms Golgi-specific lipids into Golgi-like nanotubules. *J. Biol. Chem.* **276**, 41224–41228 (2001).
91. Stahelin, R. V. *et al.* Contrasting membrane interaction mechanisms of AP180 N-terminal homology (ANTH) and epsin N-terminal homology (ENTH) domains. *J. Biol. Chem.* **278**, 28993–28999 (2003).

Acknowledgements B. Peter provided inspiration for this review, and although he has moved on to better things, his thoughts and contribution were invaluable. He is largely responsible for Fig. 3. We also thank P. Evans and all members of the laboratory for their continuous curvature discussion. J.G. was the recipient of an MRC Predoctoral Fellowship and Karn Fund Postdoctoral Fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.T.M. (hmm@mrc-lmb.cam.ac.uk).

Organelle identity and the signposts for membrane traffic

Rudy Behnia¹ & Sean Munro¹

Eukaryotic cells have systems of internal organelles to synthesize lipids and membrane proteins, to release secreted proteins, to take up nutrients and to degrade membrane-bound and internalized molecules. Proteins and lipids move from organelle to organelle using transport vesicles. The accuracy of this traffic depends upon organelles being correctly recognized. In general, organelles are identified by the activated GTPases and specific lipid species that they display. These short-lived determinants provide organelles with an identity that is both unique and flexible. Recent studies have helped to establish how cells maintain and restrict these determinants and explain how this system is exploited by invading pathogens.

A cell biologist peering at an electron micrograph, or a biochemist fractionating pulverized cells, is faced with the problem of distinguishing between diverse membrane-bound organelles. Proteins floating in the cytoplasm of the cell often face a similar problem: they have to specifically recognize one organelle among all the others. Many of the proteins that mediate traffic between organelles are 'peripheral' membrane proteins that lack transmembrane domains and instead are recruited directly to the cytosolic surface of the organelles on which they act. Such proteins include the various coat proteins that generate transport vesicles, the motor proteins that move vesicles and organelles along the cytoskeleton and the 'tethering factors' that attach the vesicles to their destination organelles before fusion (Fig. 1)¹.

The accurate functioning of membrane traffic requires that these peripheral proteins bind only to one specific organelle. In some cases they recognize a specific 'integral' membrane protein anchored in the bilayer of the relevant organelle. For instance, secretion is initiated when cytosolic ribosomes dock onto the receptors and the translocation machinery that are integral to the membrane of the endoplasmic reticulum (ER). Membrane traffic also involves other integral membrane proteins such as the SNAREs (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors) that mediate bilayer fusion, and the cargo receptors that collect soluble proteins from the organelle lumen into transport vesicles (Fig. 1).

However, the majority of membrane-traffic components are peripheral membrane proteins that recognize the correct organelle by binding to either specific lipids, such as phosphoinositides, or to activated forms of GTPases. These lipids and GTPases are usually present on only a subset of internal membranes and hence provide each organelle with a unique identity that allows it to be recognized by the many proteins that act on its cytosolic surface.

The reason that short-lived molecules such as phosphoinositides or activated GTPases are so widely used as internal spatial signals probably reflects the ease with which their subcellular distribution can be controlled. In contrast to integral membrane proteins, which have to be inserted into the ER and then trafficked through a series of compartments to their site of action, lipids can be synthesized, or GTPases activated, only at the site at which they are required. This confers the advantage of accuracy, but also plasticity, allowing vesicles to rapidly lose the identity of the organelle from which they have budded. In

addition, it is believed that some organelles themselves can change their identity. Examples include the early endosomes or the cisternae of the Golgi, which continuously mature into later compartments as new ones are generated to replace them.

Although this plasticity has enabled internal organelles to function efficiently, it also seems to be an Achilles' heel for eukaryotic cells that is exploited by invading pathogens. Many pathogenic bacteria and protozoa replicate inside our cells in membrane-bound compartments where they can hide from immune surveillance. The pathogens form and maintain these compartments by altering organelle identity and subverting normal membrane-traffic events. This alteration can be achieved by simply perturbing cytoplasmic GTPases and lipid species, and it does not require the removal or addition of integral membrane proteins. Thus the plasticity of organelle identity may have left eukaryotic cells particularly vulnerable to invasion.

For these reasons the regulation of organelle-specific GTPases and lipids has emerged as an issue relevant to both fundamental cell biology and medicine. In this review, we summarize what is known about how GTPases are activated and lipids are synthesized to create organelle identity. We also mention recent work suggesting that transient attachment of fatty acids contributes to the recruitment of peripheral membrane proteins. We finish by describing how these systems are subverted by invading pathogens.

Organelle-specific GTPases

Two main classes of small GTPase, the Rab and Arf families, contribute to defining the identity of organelles². These GTPases are molecular switches that can alternate between a GTP-bound active form and a GDP-bound inactive form³. The latter is cytosolic, whereas the active form is associated with membranes. These key characteristics enable them to recruit peripheral membrane proteins only when associated with specific membranes. The overall role of the Rab and Arf GTPases in defining organelle identity is similar, but they differ in the way that they achieve their membrane targeting, and so we will discuss them separately.

Rab GTPases

There are more than 60 different Rab proteins in mammals. Some of these are ubiquitously expressed whereas others are cell-type specific

¹MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

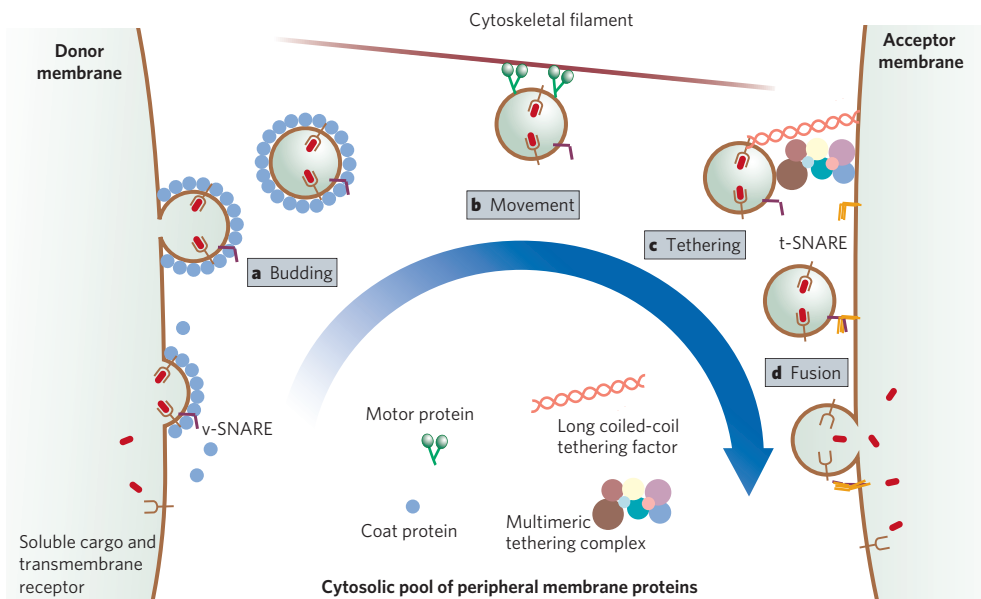


Figure 1 | Schematic representation of the steps of vesicle transport. **a**, Coat proteins are recruited to the cytosolic face of the donor membrane and induce the formation of a vesicle. The coat recruits SNAREs and transmembrane receptors bound to their cargo. **b**, After uncoating, motor protein can be recruited to enable the vesicle to travel along microtubules or actin filaments. **c**, Once at its destination, the vesicle becomes tethered to the acceptor membrane, probably by long coiled-coil proteins or multimeric tethering complexes. **d**, The SNAREs on the vesicle and acceptor membrane form a complex which drives membrane fusion and hence delivery of the contents of the vesicle.

and are associated with specialized organelles⁴. There are fewer Rabs in lower eukaryotes, which suggests that additional specialist Rabs emerged as new cell types arose during evolution. All Rabs so far characterized are associated with specific organelles, where they act as targeting determinants for a wide variety of protein, including molecular motors and tethering factors^{2,5}. For example, the ubiquitous Rab6 localizes to the *trans*-Golgi where it recruits Bicaudal D, an accessory protein for the microtubule motor dynein^{6,7}, and TMF1, a coiled-coil protein involved in membrane traffic⁸. Another ubiquitous Rab, Rab5, recruits many proteins to early endosomes including EEA1, a tethering factor for endosome fusion⁹. By contrast, Rab27a is expressed only in secretory cells such as melanocytes and recruits the motor myosin-Va to the membranes of melanosomes where it mediates anchoring to the actin cytoskeleton¹⁰.

Rab GTPases have a central position in determining when and where peripheral membrane proteins are recruited to organelles. But how are Rab GTPases directed to the correct membrane? Rabs are modified by the attachment of prenyl groups, usually two of them, to the carboxy terminus of the protein. This lipid anchor, although essential for membrane association, is clearly not enough to determine the subcellular localization of these GTPases because it is shared by many different Rabs^{11,12}. The linker between the lipid anchor and the rest of the protein, the 'hypervariable region', shows the highest level of sequence divergence between Rabs and have been proposed to act as a signal for targeting¹³. However, recent analysis of several Rabs has shown that this region is neither necessary nor sufficient for correct targeting, and so it does not seem to be a general signal for Rab targeting¹⁴. Instead, current evidence suggests that the correct recruitment of Rab GTPases is governed by the proteins that regulate their GDP/GTP cycle (Fig. 2a), because some of these are themselves found only on specific organelles.

Rab proteins bound to GDP form a complex in the cytosol with a protein called the GDP-dissociation inhibitor (GDI). This is thought to mask the prenyl groups and so prevent random association with membranes. A set of membrane proteins known as GDI displacement factors (GDFs) are thought to catalyse the dissociation of Rab from GDI at the target membrane, resulting in the anchoring of the prenyl groups in the lipid bilayer^{11,12}. Once on the membrane, the Rab is activated by replacement of the GDP by GTP by a guanine nucleotide exchange factor (GEF). As with other Ras superfamily GTPases, GTP binding induces a conformational change confined to two main segments, called the switch 1 and switch 2 regions, contributing to the creation of an interface for effector proteins to bind specifically to the active GTP-bound conformation. The Rabs themselves hydrolyse GTP

very inefficiently, and, *in vivo*, this process is stimulated by GTPase-activating proteins (GAPs), after which free GDI can then extract the inactive Rab from the membrane.

According to this cycle, GDFs and GEFs are the main determinants for the localization of active Rabs. However, the relative importance of these two classes of protein for specificity is unclear. Several small membrane proteins from the well-conserved Yip1, Yip2/Yop1 or Pra1/Yip3 families bind promiscuously *in vitro* to prenylated Rabs and have been proposed to act as GDFs^{15,16}. One of these, Pra1, catalyses the displacement of Rab9 from GDI *in vitro*^{17,18}. However, the Yip1, Yip2 and Pra1 families have many fewer members than there are Rabs, which suggests that these proteins are, at most, only partly responsible for the specific targeting of Rabs. Indeed, all eight members of the human Yip1 family are located at the ER and Golgi¹⁹. Moreover, if a GDF dissociates a Rab from GDI on a membrane where the relevant GEF is absent then the GDP-bound Rab would be rapidly removed from the membrane by GDI. Thus the importance of the Yip1 and Pra1 families for specificity *in vivo* remains to be proven. Interestingly, both Yip1 and Yip3 (the yeast Pra1 homologue) have a role in ER-to-Golgi transport that is independent of their interaction with Rabs, suggesting that the families may have other functions^{20–22}.

It seems likely that the Rab GEFs make at least some contribution to the organelle-specific recruitment of Rabs. However, despite the large number of Rabs, only a few Rab GEFs have so far been identified, and thus much remains unclear about their role^{11,12}. Nonetheless, the known GEFs seem to be localized to specific organelles, implying that their distribution could determine the distribution of the relevant Rab. Moreover, all known GEFs are peripheral-membrane proteins, and studies of how they are themselves localized are beginning to suggest general principles for organization of internal membranes^{23,24}. An interesting case is that of the Rab protein Sec4, which is present on the *trans*-Golgi and post-Golgi vesicles in yeast. The GEF for this GTPase, the peripheral membrane protein Sec2, is an effector of another Rab GTPase called Ypt31 (ref. 25). Similarly, it has been suggested that the GEF for Ypt31 is recruited on membranes by yet another Rab, Ypt1, localized on the *cis*-Golgi²⁶. This 'Rab cascade' could be a way to coordinate the sequential generation of spatial landmarks if, as is widely believed, the Golgi cisternae mature from *cis* to *trans* through the stack. Another illuminating case is mammalian Rab5 on early endosomes, which seems to be initially recruited by a GEF in the nascent clathrin-coated vesicles, which once budded will fuse to form the early endosome²⁷. However, Rab5 itself binds directly to a complex containing Rabex-5, a second Rab5 exchange factor²⁸. Recruitment by the GTPase of its own GEF clearly provides an amplification mechanism

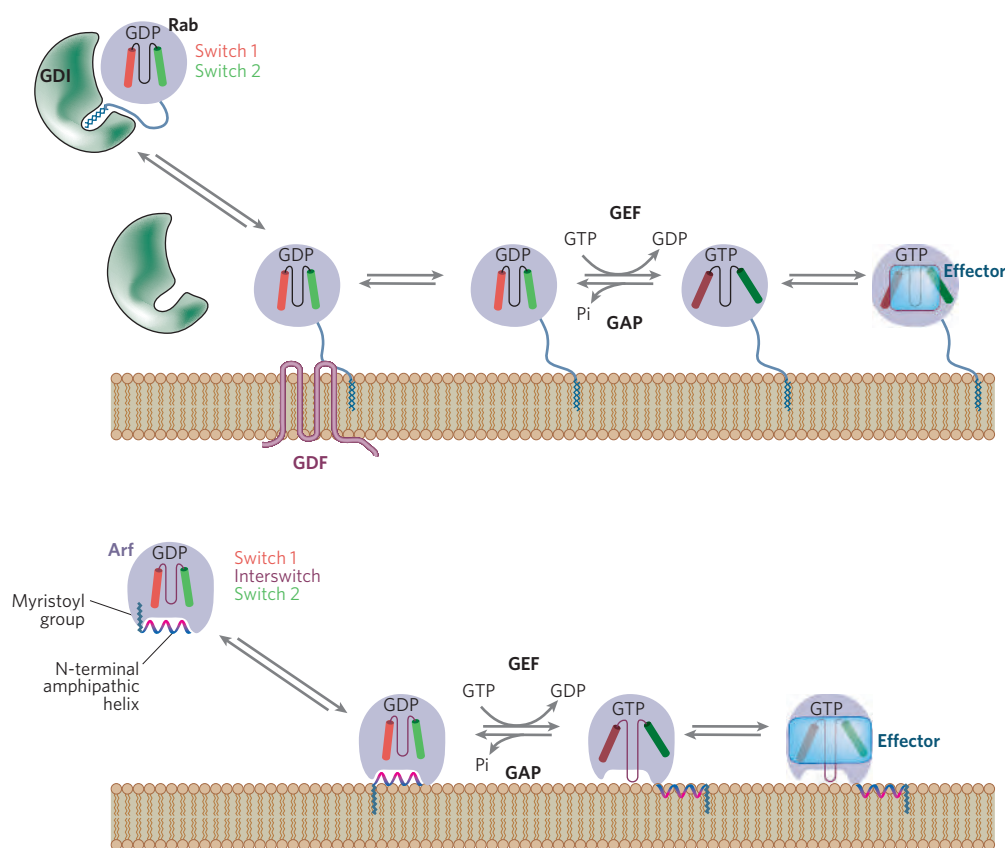


Figure 2 | Recruitment of the Rab and Arf family GTPases to membranes.
a, Rab GTPases. Rab-GDP forms a complex in the cytosol with GDI. GDF displaces GDI from Rab-GDP, and the Rab is anchored at the membrane by its C-terminal prenyl groups. There, a GEF activates the Rab by exchange of GDP for GTP. This induces a conformational change in the switch 1 and 2 regions of the GTPase (indicated here by a change in colour) that enables the Rab-GTP to bind to its effectors. A GAP stimulates the hydrolysis of GTP and the Rab is retrieved by GDI to the cytosol once again (not shown).
b, Arf GTPases. In the cytoplasm, the N-terminal amphipathic helix of Arf-GDP is tucked into a hydrophobic pocket. The N-terminal myristoyl group binds reversibly to membranes where a GEF activates the Arf. The exchange of GDP for GTP induces a change not only in switch 1 and 2 but also in the interswitch loop which moves to displace the N-terminal helix out of its pocket. Arf-GTP then binds tightly to membranes through the hydrophobic residues of the N-terminal helix and the myristoyl anchor, and recruits effectors. A GAP reverses the process and returns the Arf to the cytosol.

that could sustain recruitment of Rab5 to membranes once endosomes have formed.

Finally, it is worth noting the contribution of the Rab GAPs, which exert a negative effect on Rab distribution. A number of these have been identified, and their biological importance for membrane traffic has been confirmed by studies in yeast and mammalian cells^{29–31}. However, their precise contribution to Rab location, and how their distribution is regulated, are at present poorly understood.

ARF GTPases

The Arf family of GTPases comprises Sar1, Arf1–6 and a number of Arf-like GTPases that are similar to Arfs but more distantly related. In common with the Rabs, the Arfs are localized to specific organelles in the cell and recruit a wide range of effectors (Fig. 3). Sar1 recruits the COPII vesicle coat to the membranes of the ER³², whereas Arf1 recruits COPI and clathrin/adaptor vesicle coats to the Golgi and also other effectors including putative vesicle tethers^{33,34}. Arf6 is localized to the plasma membrane, where it is responsible for the recruitment of the kinase that makes phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂) (see below)³⁵. The roles of Sar1 and Arf1 are well established, but the localization and function of most of the Arf-like proteins are less well characterized. Exceptions are Arl1 and ARFRP1 (Arl3 in yeast), which act in a pathway involved in the recruitment to the *trans*-Golgi of a set of coiled-coil proteins that share a C-terminal GRIP domain that is directly bound by Arl1-GTP^{36,37}.

Like Rabs, Arf GTPases carry a lipid moiety, which in the case of Arfs is an amino-terminal myristoyl group. Again, this modification is not sufficient for specific targeting as it binds promiscuously to membranes. Moreover, a few members of this family such as Sar1 or Arl3/ARFRP1 lack this modification³⁸. The specificity of the interaction of Arfs with membranes must therefore be conferred by something else, and in this case it primarily seems to be at the level of the GDP/GTP switch catalysed by their GEFs.

Arf GTPases have the peculiarity of possessing an N-terminal amphipathic helix that follows the myristoyl lipid group in the struc-

ture. In the GDP-bound conformation, the hydrophobic residues of this amphipathic helix are masked inside a hydrophobic pocket in the core of the GTPase³⁸. The myristoyl group mediates rapidly reversible and non-specific association with membranes. When bound to membranes, the exchange of GDP for GTP is catalysed by a specific GEF (Fig. 2b). As with Rabs, this exchange induces a conformational change in the GTPase. This involves not only the classic switch 1 and switch 2 regions, but also an intermediate region specific to Arfs termed the 'interswitch' by Pasqualato *et al.*³⁸. In the GTP-bound form, this interswitch slides through the protein and pushes the N-terminal amphipathic helix out of its pocket and forces it to interact with an adjacent bilayer (Fig. 2b)^{39,40}. This mechanism allows the direct coupling of GTP-binding to membrane recruitment, and thereby explains the absence of an equivalent to Rab GDI in the Arf cycle. In the GTP-bound form, Arf family GTPases bind to their specific effectors, until a GAP induces the hydrolysis of GTP, which reverses membrane association and effector binding.

This structural evidence that the activation of Arfs is coupled with their membrane association indicates that their distribution is also controlled by that of the GEFs that activate them. A simple example is that of Sar1, which is recruited to ER membranes upon activation by the GEF Sec12, a transmembrane protein embedded in the ER³². But in the case of Arfs, all GEFs identified so far are peripheral membrane proteins that contain a 200-residue conserved 'Sec7' catalytic domain⁴¹. These GEFs are recruited to either the Golgi or the plasma membrane. In most cases the mechanism of their recruitment is not fully understood but is likely to involve the contribution of several lipid and/or protein interactions that together define a compartment or a subcompartment. For example, Arf1 is activated by a GEF called GBF1 on the *cis*-Golgi. GBF1 interacts with the peripheral protein p115 (ref. 42) and also with a transmembrane protein called Gmh1 (ref. 43). Gea2, a yeast homologue of GBF1, also binds to the P-type ATPase Drs2, integral to membranes⁴⁴. However, none of these interactions by themselves seems to be essential to localize Arf1 to the *cis*-Golgi. The precise mechanism by which GBF1 and its relatives on the *trans*-Golgi (BIG1 and BIG2) are

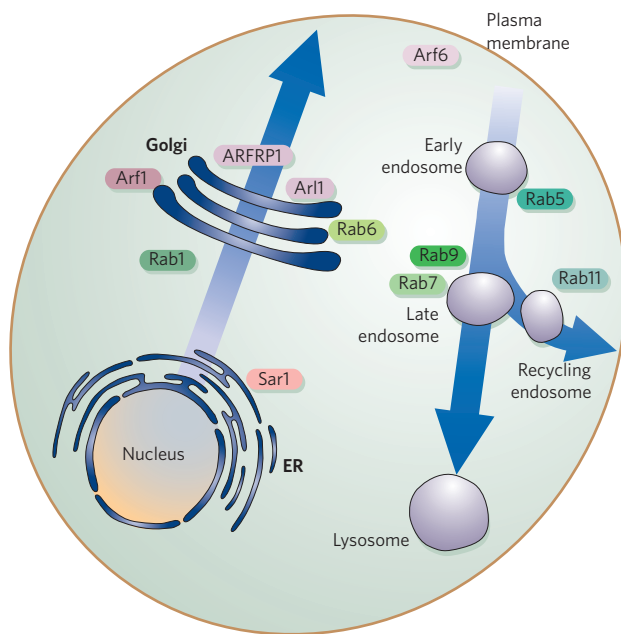


Figure 3 | Location of GTPases to specific organelles. Examples of GTPases that are localized to specific membranes within eukaryotic cells. Arf family GTPases are shown in purples, and Rabs in greens. Some, such as Arf1, are found in multiple locations (in this case throughout the Golgi stack), but most are restricted to only one organelle. The examples shown are the best-characterized and ubiquitous cases, but there are many more GTPases, especially Rabs, that are less well characterized or are found in specialized or polarized cells. A detailed account of the distribution of all Rabs characterized so far can be found elsewhere^{2,102}.

recruited remains somewhat mysterious. At the plasma membrane, the levels of Arf6-GTP are regulated in response to external signals. Arf GEFs of the cytohesin/ARNO family contain pleckstrin homology domains that mediate binding to phosphoinositides generated at the plasma membrane by signal-transduction pathways⁴⁵.

Transmembrane proteins may also contribute to the recruitment of several members of the Arf family. The GDP-bound form of Arf1 has been reported to bind to the cytoplasmic domain of a member of the p24 family of cargo receptors that cycles between the ER and Golgi⁴⁶. More recently, Honda *et al.* have suggested that a SNARE called membrin could be a receptor for Arf1 on the *cis*-Golgi⁴⁷. Finally, yeast Arl3 and its human counterpart ARFRP1 lack a myristoylation site. They are instead N-terminally acetylated, and this modification is necessary for their binding to a small polytopic membrane protein Sys1 that acts as a receptor on the membranes of the *trans*-Golgi^{48,49}.

In common with the Rab cycle, GAPs play a key role in the Arf cycle and seem to be particularly important for removing the active GTPases from the membrane leaving a compartment in vesicles. A striking example is that of Sec23, the GAP for Sar1, which is a component of the COPII coat recruited by Sar1 itself³². Sec23 induces Sar1 to hydrolyse GTP after vesicle budding, leading to vesicle uncoating. In the case of COPI, the coat itself is not a GAP; instead ArfGAP1 binds to the coat so as to be recruited into the budding vesicle⁵⁰. In addition, ArfGAP1 has the unusual property of being highly sensitive to the degree of membrane curvature: the rate of ArfGAP1-stimulated hydrolysis of GTP on Arf1 increases with the curvature of the lipid bilayer⁵¹ (see also the review by McMahon and Gallop, p.590, in this issue). This leads to a model in which polymerization of the COPI coat increases membrane curvature, which in turn increases the rate of hydrolysis of GTP on Arf1, coupling budding to removal of Arf1 and hence coat disassembly.

Organelle-specific phosphoinositides

The second major class of molecule that contributes to the unique identity of organelles is specific lipid species, especially phosphoinosi-

tides. These are forms of phosphatidylinositol (PtdIns) with phosphate attached by specific kinases to the 3, 4 or 5 positions of the inositol ring (Fig. 4a). Two of the phosphoinositides are second messengers and are synthesized only in response to external signals (PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃). However, the majority are constitutively present in cells and are generally found on only one or a small subset of organelles (Fig. 4b). They are recognized by peripheral-membrane proteins that usually bind specifically to that particular phosphoinositide^{52,53}.

Perhaps the best characterized is PtdIns(3)P, which is present on early endosomes, and is recognized by a wide range of peripheral membrane proteins that have key roles in endosomal function. These include motor proteins, tethering proteins and the machinery that downregulates receptors by means of membrane invagination and the formation of multivesicular bodies. Many of these proteins recognize PtdIns(3)P through one of two small folds: the FYVE domain and the PX domain^{54,55}. PtdIns(3)P is also used as a substrate for the synthesis of PtdIns(3,5)P₂, which is found on later endocytic compartments. Although the importance of PtdIns(3,5)P₂ for late endosomal function is clear from genetic studies in yeast, its effectors have been more elusive. Indeed it is not clear if it is localized to late endosomes, lysosomes or both⁵⁶. However, two proteins have recently been identified that bind specifically to PtdIns(3,5)P₂, including a coated vesicle component, and so these issues should soon be clarified^{57,58}.

Phosphoinositides are also important in the exocytic pathway, with PtdIns(4)P present on the Golgi, where it is recognized by vesicle coat proteins and by the pleckstrin homology domains of proteins that deliver lipids to the Golgi^{59–61}. PtdIns(4)P is also found on the plasma membrane, where it is a substrate for the synthesis of PtdIns(4,5)P₂. The proteins that are specifically localized to the *trans*-Golgi and bind PtdIns(4)P seem to also recognize the Golgi GTPase Arf1 (ref. 62). This combinatorial recognition of two determinants allows more restricted targeting than would be seen with either alone, and it is emerging as a more general feature of targeting of peripheral membrane proteins.

PtdIns(4)P and PtdIns(4,5)P₂ at the plasma membrane can be phosphorylated by 3-kinases to generate signalling lipids, and PtdIns(4,5)P₂ can also be cleaved to generate the second messengers diacylglycerol (DAG) and Ins(1,4,5)P₃. However, PtdIns(4,5)P₂ is also a major landmark for proteins that need to find the plasma membrane. As such, it is recognized by several key proteins involved in the formation of clathrin-coated pits, including the AP2 adaptor, as well as by many proteins that regulate the actin cytoskeleton^{63,64} (for a full discussion of PtdIns(4,5)P₂ see the review by McLaughlin and Murray, p. 605 in this issue). PtdIns(4,5)P₂ has also been proposed to cluster to define subdomains within the plasma membrane⁶⁵, but this has recently been challenged by detailed *in vivo* imaging⁶⁶.

The role of phosphoinositides as key landmarks for subcellular organization raises the obvious question of how their restricted distribution is established. Analogous to the GTPases described above, the answer is that this is achieved by a combination of localized synthesis by specific kinases and rapid turnover by phosphatases that prevent the lipid spreading between compartments. Strikingly, all phosphoinositide (PI) kinases are themselves peripheral membrane proteins, and, at least in some cases, their location is determined by organelle-specific GTPases^{23,24}. For instance, the endosomal PtdIns-3-OH kinase (PI(3)K) Vps34 that makes PtdIns(3)P is regulated, in part, by the endosomal GTPase Rab5 (ref. 67), whereas the PI(5)K that makes PtdIns(4,5)P₂ is recruited to the plasma membrane by Arf6 (ref. 35). In addition, organelle-specific PI kinases seem to be major targets for regulatory pathways that adjust the constitutive level of lipid to suit the changing needs of the cell. Thus the endosomal PI(3)K Vps34 is part of a large complex that contains a protein kinase and additional components that seem to recruit the kinase to membrane structures formed during autophagy⁶⁸. Likewise, the PI(5)K that generates PtdIns(3,5)P₂ on later endocytic compartments interacts with further proteins and is apparently regulated by osmotic stress⁵⁶.

In addition to the kinases, phosphatases are important for restrict-

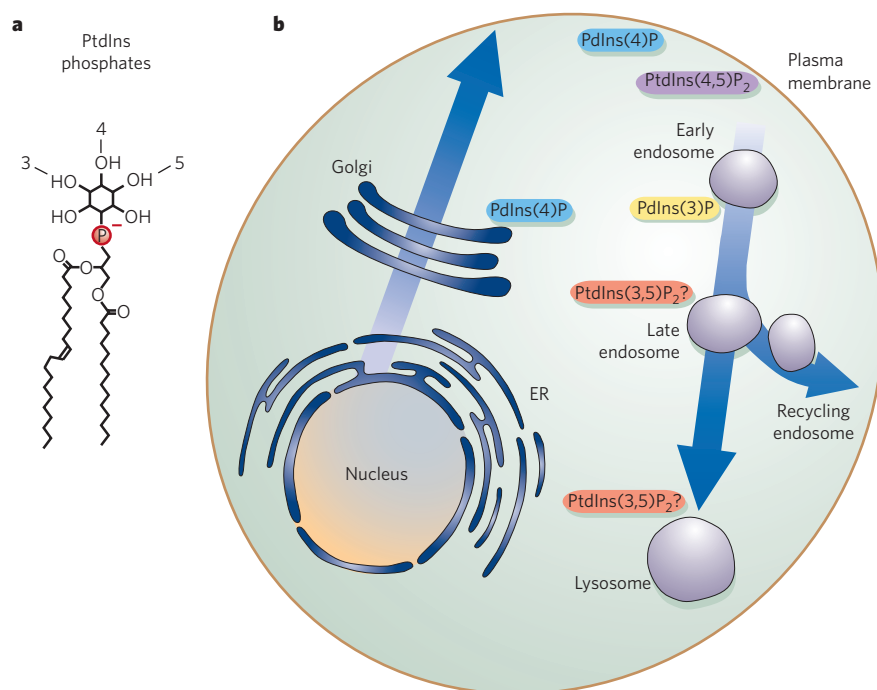


Figure 4 | Location of phosphoinositides to specific organelles. **a**, Phosphoinositides are generated by phosphorylation of phosphatidylinositol (PtdIns) on the 3, 4 or 5 positions of the inositol ring. All seven possible combinations of phosphorylation occur *in vivo*. PtdIns is shown with a $C_{18:1}$ unsaturated chain, but others can also occur *in vivo* especially $C_{20:4}$ arachidonate. **b**, A schematic illustration of the major sites of intracellular accumulation of the well-characterized phosphoinositides. This summary is primarily based on studies in yeast and mammalian tissue culture cells and it remains to be seen how much variation occurs in specialized cells. PtdIns(3,5)P₂ has not been localized in mammalian cells, but the 5-kinase is found on late endosomes, and in yeast the lipid accumulates on the vacuole⁵⁶. Further discussion of phosphoinositide distribution can be found in recent reviews^{52,53}.

ing the distribution of particular lipids. For instance, the phosphatase synaptojanin is recruited to clathrin-coated pits at the plasma membrane and seems to be important for preventing PtdIns(4,5)P₂ from entering the endocytic pathway^{53,69}. Another example is the PI 4-phosphatase Sac1, which is localized in the ER in both yeast and mammals and prevents PtdIns(4)P from accumulating on ER membranes^{70,71}. In addition, yeast Sac1 relocates to the Golgi when nutrients are limiting, where it reduces the level of this PtdIns(4)P and hence the rate of secretion^{70,71}.

Finally, it should be noted that the spatial regulation of phosphoinositides is complicated by their metabolic inter-relationship. Thus, some phosphoinositides must exist before others can be generated. Degradation of one can also generate another. This means that the same lipid can be generated by multiple routes; for instance, PtdIns(3)P on endosomes or phagosomes can be generated both by the action of a PI(3)K on PtdIns or by dephosphorylation of the signalling lipids PtdIns(3,4)P₂ and PtdIns(3,4,5)P₃ coming in from the plasma membrane⁷². Moreover, it is striking that the basic exocytic routes and endocytic routes are characterized by monophosphorylated lipids (PtdIns(3)P and PtdIns(4)P) being elaborated further by 5-phosphorylation, which might reflect an underlying organizing principle of membrane traffic that emerged early in evolution⁵³.

Further roles for lipids in protein targeting

In the context of lipids and organelle identity it is also worth mentioning the role of DAG and phosphatidylserine (PtdSer). DAG is generated during signal-transduction events at the plasma membrane and transiently recruits a number of signalling proteins, most notably protein kinases C. However, it is also constitutively produced by lipid metabolism on Golgi membranes and seems to be recognized by a small number of Golgi proteins, although how these proteins distinguish Golgi and plasma membrane pools of DAG remains unclear⁷³.

PtdSer is an acidic phospholipid found throughout the membranes of the cell. It is normally present in both leaflets of the bilayer, except in the plasma membrane, and perhaps in endocytic compartments. In these membranes it is asymmetrically distributed as it is all translocated to the leaflet facing the cytoplasm. It seems that the resulting twofold increase in PtdSer levels in the cytoplasmic leaflet is important for a number of peripheral proteins to be specifically recruited to the plasma membrane through electrostatic interactions with stretches of basic residues⁷⁴. The asymmetric distribution of PtdSer requires spe-

cific translocases, and, interestingly, a number of putative translocases are required for membrane traffic events at the plasma membrane or endosomes^{75,76}.

Finally, an exciting area of study to have emerged recently is that of the direct covalent attachment of proteins to lipids as a mechanism to recruit proteins to particular organelles. One unusual but striking example is Atg8, a key regulator of autophagy, which is recruited to autophagosomes by reversible attachment to phosphatidylethanolamine⁷⁷. However, perhaps the most widespread example is palmitoylation. Many peripheral membrane proteins are modified by the addition of one or more palmitoyl ($C_{16:0}$) fatty acyl groups that contribute to the attachment of the protein to the cytosolic face of the membrane⁷⁸. Analysis of the biology of palmitoylation was hampered until recently by the lack of identified protein palmitoyltransferases (PATs). The recent identification of two PATs in yeast, and the realization that they share a 'DHHC' motif found in a larger family of membrane proteins of unknown function, suggest that palmitoylation may be more specific than previously suspected⁷⁹. This means that the distribution of particular PATs could control the distribution of their substrates. Thus PATs in the Golgi seem to attach a number of proteins such as Ras and nitric oxide synthase to Golgi membranes⁷⁹. The palmitoylated proteins can then travel on post-Golgi vesicles to the plasma membrane. In addition, recent studies have shown that proteins can be depalmitoylated and hence released from the membrane to travel to new locations^{80,81}.

Identity theft

Many pathogenic bacteria invade host cells and in the process form membrane-bound compartments (similar to the cellular organelles) in which they replicate. These bacteria have developed a number of tactics to disguise these compartments so that they seem to be benign parts of the cell. This allows them to avoid the cell's defence mechanisms and instead receive vesicles delivering lipids and nutrients. Bacteria steal the identity of cellular organelles by manipulating the lipids and GTPases that define organelle identity^{82,83}.

Bacterial pathogens enter mammalian cells by two main pathways⁸⁴. One route takes advantage of professional phagocytic cells such as macrophages that take up microorganisms into phagosomes. In a second route, bacteria force their way inside non-phagocytic cells by using 'secretion systems' to inject virulence factors across the plasma membrane of the host cell⁸⁵. These virulence proteins trigger localized

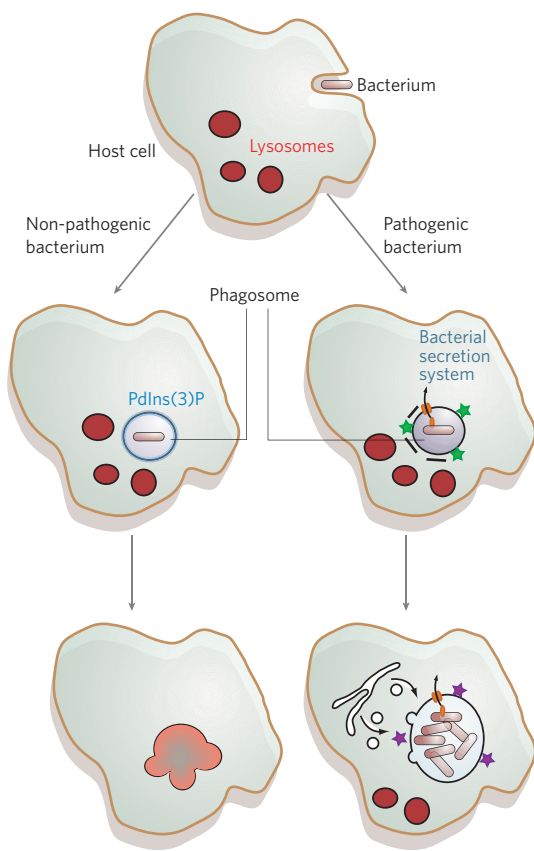


Figure 5 | Identity theft by invading pathogens. Bacteria are taken up by phagocytic cells actively seeking to destroy invaders or can force non-phagocytic cells to engulf them. In either case they end up in a phagosome — an intracellular membrane-bound compartment that has many of the characteristics of an early endosome. This compartment matures and fuses with lysosomes, leading to the digestion of the bacteria (left). However, many pathogenic bacteria can evade this fate by using their secretion systems to transfer virulence factors into the cytosol of the host cells that perturb the pathways that generate organelle identity (right). These secreted virulence factors arrest the maturation of the compartment (green stars), and once the endocytic pathway has been evaded, further proteins are secreted (purple stars) to set up interactions with the secretory pathway to attract the delivery of host-cell proteins and lipids while the bacteria replicate.

alteration of the cytoskeleton and the engulfment of the bacteria. For example, *Salmonella* induces extensive actin rearrangements and membrane ruffling at the site of bacteria–host-cell contact. The bacteria inject SigD/SopB, a phosphatidylinositol phosphatase, to degrade the plasma-membrane lipid PtdIns(4,5)P₂ that normally recruits proteins that organize cortical actin⁸⁶.

Irrespective of the route of entry, bacteria would normally follow the endocytic pathway, which consists of the maturation of the membrane-bound compartment to a late endosome state followed by fusion with lysosomes, leading to digestion of the trapped microbe (Fig. 5). But many pathogens have developed strategies to avoid this fate. Some bacteria such as *Listeria* and *Shigella* simply lyse their plasma-membrane-derived compartment and escape into the cytoplasm⁸⁷. However, most species remain in their compartment, or vacuole, and prevent its maturation to avoid fusion with lysosomes (Fig. 5). For example, compartments containing *Mycobacterium tuberculosis* retain determinants of early endosomal organelles such as Rab5 and exclude the late endosomal determinant Rab7 (ref. 88). In addition, mycobacteria secrete an enzyme SapM that dephosphorylates the PtdIns(3)P on its compartment, thereby inhibiting movement down the endocytic route toward fusion with lysosomes⁸⁹. By contrast, *Salmonella* arrests maturation at a later stage. It does not remove PtdIns(3)P from the membrane of its compartment and so allows fusion with early endo-

somes to increase compartmental size. Instead, the secreted bacterial phosphatase SopB is necessary for the generation and persistence of this PtdIns(3)P, perhaps by degrading PtdIns(3,5)P₂ and arresting the progression of the compartment towards fusion with lysosomes⁹⁰.

Once they have evaded destruction by lysosomal hydrolases, the invading bacteria then set up interactions with the secretory pathway to acquire nutrients (Fig. 5). For example, vacuoles containing *Legionella* or *Brucella* intercept ER-derived vesicles⁹¹. *Legionella* secretes a protein called RalF out of its vacuole, which contains a domain closely related to the Sec7 domain of the GEFs for Arf1 (ref. 92). RalF is sufficient to recruit Arf1 to the membrane of the vacuole and thus perhaps mimic the vesicle-docking machinery on the *cis*-Golgi. However, RalF is not essential for bacterial replication and redirection of ER-derived material, and so other, as yet unidentified, bacterial proteins must be involved. *Chlamydia*- and *Salmonella*-containing vacuoles reach a perinuclear position in the host cell by movement along microtubules, where they recruit different Rabs and intercept vesicles carrying proteins and lipids trafficking from the Golgi apparatus^{93,94}. The positioning and the structure of such vacuoles are the result of a fine-tuning between the recruitment of both the plus-end motor kinesin and the minus-end motor dynein on their membrane mediated in part by Rab7 (refs 94–96).

It thus seems that bacteria use a combination of strategies to perturb, both spatially and temporally, the pathways that generate membrane identity. This enables them to recruit many host-cell peripheral membrane proteins and use them for their own malevolent purposes. Interestingly, there is recent evidence that viruses can use similar strategies. For example, vaccinia virus recruits motor proteins to its external membrane to travel to the plasma membrane⁹⁷, and poliovirus recruits Arf GTPases to its replication sites⁹⁸. This field has made rapid progress, but the fact that many of the bacterial proteins secreted across vacuolar membranes do not yet have assigned functions indicates that there remains much to be learnt.

Future questions

From the work of many laboratories it is now clear that the organization of the internal membranes of eukaryotic cells is dependent on the generation of spatially restricted GTPases and lipids. Nonetheless, there still remains much to be learnt about organelle identity. The inherent appeal of simplicity makes the notion of a GTPase or lipid precisely restricted to one organelle very attractive. However, living cells are rarely simple, and there are clear cases where determinants are found in more than one place, such as PtdIns(4)P on the *trans*-Golgi and the plasma membrane, or Arf1 throughout the Golgi stack. In some cases this may allow specialized sets of proteins to be recruited to multiple locations, or it may simply reflect that complete precision requires the combinatorial recognition of multiple determinants. For any given protein, it may therefore be important to search for additional binding partners even after one has been discovered. Indeed, further factors such as membrane shape could also contribute.

Nonetheless, most GTPases and phosphoinositides have a restricted distribution and there must be mechanisms that account for this. It is clear that there are still a number of missing components, in particular the GEFs for many members of the Rab and Arf families. However, it is apparent from the GTPase regulators and phosphoinositide enzymes so far identified that most are themselves peripheral membrane proteins and hence must recognize organelle-specific determinants. Thus, perhaps the major challenge is to understand what ultimately initiates these pathways that lead to the generation of localized determinants for recruiting effectors.

The impossibility of organelle identity being defined by an infinite series of peripheral proteins raises the issue of the importance of integral membrane proteins. The fact they must start their lives in the ER makes them seem a liability as spatial determinants. Indeed it is striking that the ER itself is an exception as an organelle because most of the known peripheral membrane proteins bind directly to integral membrane proteins⁹⁹, with just one GTPase (Sar1) and no specific lipids

found on this organelle. The SNAREs are integral membrane proteins found in many organelles that have been proposed to contribute to the specificity of membrane fusion, although this has been questioned by their promiscuous assembly *in vitro*^{100,101}. Determining whether the SNAREs contribute to the specificity of membrane traffic, and how they interact with the peripheral proteins, is a key question for the future. Nonetheless, other integral membrane proteins could contribute elsewhere if they were activated when in the right location by changes in luminal pH, Ca²⁺ or levels of lipids such as cholesterol and sphingolipids, which might reflect progress through the secretory or endocytic routes. Alternatively, integral membrane proteins could be recognized in combination with other determinants, including other integral proteins, if they converge on the same organelle by cycling through a different itinerary. It is even possible that organelles are themselves sensitive to their location in the cell.

There are many areas that still need to be investigated and debated. However, it is already clear that as more is learnt about organelle identity, we can also hope to understand better the action of the microbial virulence factors that subvert identity, and perhaps one day even use this to design therapeutic strategies to combat these pathogens. ■

- Bonifacio, J. S. & Glick, B. S. The mechanisms of vesicle budding and fusion. *Cell* **116**, 153–166 (2004).
- Zerial, M. & McBride, H. Rab proteins as membrane organizers. *Nature Rev. Mol. Cell Biol.* **2**, 107–117 (2001).
- Vetter, I. R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299–1304 (2001).
- Pereira-Leal, J. B. & Seabra, M. C. Evolution of the Rab family of small GTP-binding proteins. *J. Mol. Biol.* **313**, 889–901 (2001).
- Pfeffer, S. R. Rab GTPases: specifying and deciphering organelle identity and function. *Trends Cell Biol.* **11**, 487–491 (2001).
- Matanis, T. et al. Bicaudal-D regulates COPI-independent Golgi-ER transport by recruiting the dynein-dynactin motor complex. *Nature Cell Biol.* **4**, 986–992 (2002).
- Short, B., Preisinger, C., Schaletzky, J., Kopajtich, R. & Barr, F. A. The Rab6 GTPase regulates recruitment of the dynactin complex to Golgi membranes. *Curr. Biol.* **12**, 1792–1795 (2002).
- Fridmann-Sirkis, Y., Siniosoglou, S. & Pelham, H. R. TMF is a golgin that binds Rab6 and influences Golgi morphology. *BMC Cell Biol.* **5**, 18 (2004).
- Simonsen, A. et al. EEA1 links PI(3)K function to Rab5 regulation of endosome fusion. *Nature* **394**, 494–498 (1998).
- Seabra, M. C. & Coudrier, E. Rab GTPases and myosin motors in organelle motility. *Traffic* **5**, 393–399 (2004).
- Pfeffer, S. & Aivazian, D. Targeting Rab GTPases to distinct membrane compartments. *Nature Rev. Mol. Cell Biol.* **5**, 886–896 (2004).
- Seabra, M. C. & Wasmeyer, C. Controlling the location and activation of Rab GTPases. *Curr. Opin. Cell Biol.* **16**, 451–457 (2004).
- Chavrier, P. et al. Hypervariable C-terminal domain of rab proteins acts as a targeting signal. *Nature* **353**, 769–772 (1991).
- Ali, B. R., Wasmeyer, C., Lamoreux, L., Strom, M. & Seabra, M. C. Multiple regions contribute to membrane targeting of Rab GTPases. *J. Cell Sci.* **117**, 6401–6412 (2004).
- Calero, M. & Collins, R. N. *Saccharomyces cerevisiae* Prp1p/Yip3p interacts with Yip1p and Rab proteins. *Biochem. Biophys. Res. Commun.* **290**, 676–681 (2002).
- Figuerroa, C., Taylor, J. & Vojtek, A. B. Prenylated Rab acceptor protein is a receptor for prenylated small GTPases. *J. Biol. Chem.* **276**, 28219–28225 (2001).
- Sivars, U., Aivazian, D. & Pfeffer, S. R. Yip3 catalyses the dissociation of endosomal Rab-GDI complexes. *Nature* **425**, 856–859 (2003).
- Hutt, D. M., Da-Silva, L. F., Chang, L. H., Prosser, D. C. & Ngsee, J. K. PRA1 inhibits the extraction of membrane-bound rab GTPase by GDI1. *J. Biol. Chem.* **275**, 18511–18519 (2000).
- Shakoori, A. et al. Identification of a five-pass transmembrane protein family localizing in the Golgi apparatus and the ER. *Biochem. Biophys. Res. Commun.* **312**, 850–857 (2003).
- Geng, J., Shin, M. E., Gilbert, P. M., Collins, R. N. & Burd, C. G. *Saccharomyces cerevisiae* Rab-GDI displacement factor ortholog Yip3p forms distinct complexes with the Ypt1 Rab GTPase and the reticulon Rtn1p. *Eukaryot. Cell* **4**, 1166–1174 (2005).
- Heidman, M., Chen, C. Z., Collins, R. N. & Barlowe, C. A role for Yip1p in COPII vesicle biogenesis. *J. Cell Biol.* **163**, 57–69 (2003).
- Barrowman, J., Wang, W., Zhang, Y. & Ferro-Novick, S. The Yip1p/Yif1p complex is required for the fusion competence of endoplasmic reticulum-derived vesicles. *J. Biol. Chem.* **278**, 19878–19884 (2003).
- Munro, S. Organelle identity and the targeting of peripheral membrane proteins. *Curr. Opin. Cell Biol.* **14**, 506–514 (2002).
- Munro, S. Organelle identity and the organization of membrane traffic. *Nature Cell Biol.* **6**, 469–472 (2004).
- Ortiz, D., Medkova, M., Walch-Solimena, C. & Novick, P. Ypt32 recruits the Sec4p guanine nucleotide exchange factor, Sec2p, to secretory vesicles; evidence for a Rab cascade in yeast. *J. Cell Biol.* **157**, 1005–1015 (2002).
- Wang, W. & Ferro-Novick, S. A Ypt32p exchange factor is a putative effector of Ypt1p. *Mol. Biol. Cell* **13**, 3336–3343 (2002).
- Sato, M. et al. *Caenorhabditis elegans* RME-6 is a novel regulator of RAB-5 at the clathrin-coated pit. *Nature Cell Biol.* **7**, 559–569 (2005).
- Lippe, R., Miaczynska, M., Rybin, V., Runge, A. & Zerial, M. Functional synergy between Rab5 effector Rabaptin-5 and exchange factor Rabex-5 when physically associated in a complex. *Mol. Biol. Cell* **12**, 2219–2228 (2001).
- Haas, A. K., Fuchs, E., Kopajtich, R. & Barr, F. A. A GTPase-activating protein controls Rab5 function in endocytic trafficking. *Nature Cell Biol.* **7**, 887–993 (2005).
- De Antoni, A., Schmitzova, J., Trepte, H. H., Gallwitz, D. & Albert, S. Significance of GTP hydrolysis in Ypt1p-regulated endoplasmic reticulum to Golgi transport revealed by the analysis of two novel Ypt1-GAPs. *J. Biol. Chem.* **277**, 41023–41031 (2002).
- Lafourcade, C., Galan, J. M., Gloor, Y., Haguenaer-Tsapir, R. & Peter, M. The GTPase-activating enzyme Gyp1p is required for recycling of internalized membrane material by inactivation of the Rab/Ypt GTPase Ypt1p. *Mol. Cell Biol.* **24**, 3815–3826 (2004).
- Lee, M. C., Miller, E. A., Goldberg, J., Orci, L. & Schekman, R. Bi-directional protein transport between the ER and Golgi. *Annu. Rev. Cell Dev. Biol.* **20**, 87–123 (2004).
- Gillingham, A. K., Tong, A. H., Boone, C. & Munro, S. The GTPase Arf1p and the ER to Golgi cargo receptor Erv14p cooperate to recruit the golgin Rud3p to the cis-Golgi. *J. Cell Biol.* **167**, 281–292 (2004).
- Donaldson, J. G., Honda, A. & Weigert, R. Multiple activities for Arf1 at the Golgi complex. *Biochim. Biophys. Acta* **1744**, 364–373 (2005).
- Krauss, M. et al. ARF6 stimulates clathrin/AP-2 recruitment to synaptic membranes by activating phosphatidylinositol phosphate kinase type Iγ. *J. Cell Biol.* **162**, 113–124 (2003).
- Setty, S. R., Shin, M. E., Yoshino, A., Marks, M. S. & Burd, C. G. Golgi recruitment of GRIP domain proteins by Arf-like GTPase 1 is regulated by Arf-like GTPase 3. *Curr. Biol.* **13**, 401–404 (2003).
- Panic, B., Whyte, J. R. & Munro, S. The ARF-like GTPases Arl1p and Arl3p act in a pathway that interacts with vesicle-tethering factors at the Golgi apparatus. *Curr. Biol.* **13**, 405–410 (2003).
- Pasqualato, S., Renault, L. & Cherfils, J. Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication. *EMBO Rep.* **3**, 1035–1041 (2002).
- Goldberg, J. Structural basis for activation of ARF GTPase: mechanisms of guanine nucleotide exchange and GTP-myristoyl switching. *Cell* **95**, 237–248 (1998).
- Robert, C. H., Cherfils, J., Mouawad, L. & Perahia, D. Integrating three views of Arf1 activation dynamics. *J. Mol. Biol.* **337**, 969–983 (2004).
- Jackson, C. L. & Casanova, J. E. Turning on ARF: the Sec7 family of guanine-nucleotide-exchange factors. *Trends Cell Biol.* **10**, 60–67 (2000).
- Garcia-Mata, R. & Sztul, E. The membrane-tethering protein p115 interacts with GBF1, an ARF guanine-nucleotide-exchange factor. *EMBO Rep.* **4**, 320–325 (2003).
- Chantalat, S. et al. A novel Golgi membrane protein is a partner of the ARF exchange factors Gea1p and Gea2p. *Mol. Biol. Cell* **14**, 2357–2371 (2003).
- Chantalat, S. et al. The Arf activator Gea2p and the P-type ATPase Drs2p interact at the Golgi in *Saccharomyces cerevisiae*. *J. Cell Sci.* **117**, 711–722 (2004).
- Donaldson, J. G. & Jackson, C. L. Regulators and effectors of the ARF GTPases. *Curr. Opin. Cell Biol.* **12**, 475–482 (2000).
- Gommel, D. U. et al. Recruitment to Golgi membranes of ADP-ribosylation factor 1 is mediated by the cytoplasmic domain of p23. *EMBO J.* **20**, 6751–6760 (2001).
- Honda, A., Al-Awar, O. S., Hay, J. C. & Donaldson, J. G. Targeting of Arf-1 to the early Golgi by membrin, an ER-Golgi SNARE. *J. Cell Biol.* **168**, 1039–1051 (2005).
- Behnia, R., Panic, B., Whyte, J. R. & Munro, S. Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p. *Nature Cell Biol.* **6**, 405–413 (2004).
- Setty, S. R., Strohlic, T. I., Tong, A. H., Boone, C. & Burd, C. G. Golgi targeting of ARF-like GTPase Arl3p requires its N^o-acetylation and the integral membrane protein Sys1p. *Nature Cell Biol.* **6**, 414–419 (2004).
- Liu, W., Duden, R., Phair, R. D. & Lippincott-Schwartz, J. ArfGAP1 dynamics and its role in COPI coat assembly on Golgi membranes of living cells. *J. Cell Biol.* **168**, 1053–1063 (2005).
- Bigay, J., Casella, J. F., Drin, G., Mesmin, B. & Antonny, B. ArfGAP1 responds to membrane curvature through the folding of a lipid packing sensor motif. *EMBO J.* **24**, 2244–2253 (2005).
- De Matteis, M. A. & Godi, A. PI-loting membrane traffic. *Nature Cell Biol.* **6**, 487–492 (2004).
- Wenk, M. R. & De Camilli, P. Protein-lipid interactions and phosphoinositide metabolism in membrane traffic: insights from vesicle recycling in nerve terminals. *Proc. Natl Acad. Sci. USA* **101**, 8262–8269 (2004).
- Stenmark, H., Aasland, R. & Driscoll, P. C. The phosphatidylinositol 3-phosphate-binding FYVE finger. *FEBS Lett.* **513**, 77–84 (2002).
- Ellson, C. D., Andrews, S., Stephens, L. R. & Hawkins, P. T. The PX domain: a new phosphoinositide-binding module. *J. Cell Sci.* **115**, 1099–1105 (2002).
- Efe, J. A., Botelho, R. J. & Emr, S. D. The Fab1 phosphatidylinositol kinase pathway in the regulation of vacuole morphology. *Curr. Opin. Cell Biol.* **17**, 402–408 (2005).
- Friant, S. et al. Ent3p is a PtdIns(3,5)P₂ effector required for protein sorting to the multivesicular body. *Dev. Cell* **5**, 499–511 (2003).
- Dove, S. K. et al. Svp1p defines a family of phosphatidylinositol 3,5-bisphosphate effectors. *EMBO J.* **23**, 1922–1933 (2004).
- Levine, T. P. & Munro, S. Targeting of Golgi-specific pleckstrin homology domains involves both PtdIns 4-kinase-dependent and -independent components. *Curr. Biol.* **12**, 695–704 (2002).
- Wang, Y. J. et al. Phosphatidylinositol 4 phosphate regulates targeting of clathrin adaptor AP-1 complexes to the Golgi. *Cell* **114**, 299–310 (2003).
- Balla, A., Tsyumetova, G., Tsiomenko, A., Varnai, P. & Balla, T. A plasma membrane pool of phosphatidylinositol 4-phosphate is generated by phosphatidylinositol 4-kinase type-III alpha: studies with the PH domains of the oxysterol binding protein and FAPP1. *Mol. Biol. Cell* **16**, 1282–1295 (2005).
- Shin, H. W. & Nakayama, K. Dual control of membrane targeting by PtdIns(4)P and ARF. *Trends Biochem. Sci.* **29**, 513–515 (2004).
- Honing, S. et al. Phosphatidylinositol-(4,5)-bisphosphate regulates sorting signal recognition by the clathrin-associated adaptor complex AP2. *Mol. Cell* **18**, 519–531 (2005).
- Yin, H. L. & Janmey, P. A. Phosphoinositide regulation of the actin cytoskeleton. *Annu. Rev. Physiol.* **65**, 761–789 (2003).
- Caroni, P. Actin cytoskeleton regulation through modulation of PI(4,5)P₂ rafts. *EMBO J.* **20**, 4332–4336 (2001).
- van Rhee, J., Achame, E. M., Janssen, H., Calafat, J. & Jalink, K. PIP₂ signaling in lipid domains: a critical re-evaluation. *EMBO J.* **24**, 1664–1673 (2005).

67. Murray, J. T., Panaretou, C., Stenmark, H., Miaczynska, M. & Backer, J. M. Role of Rab5 in the recruitment of hVps34/p150 to the early endosome. *Traffic* **3**, 416–427 (2002).
68. Kihara, A., Noda, T., Ishihara, N. & Ohsumi, Y. Two distinct Vps34 phosphatidylinositol 3-kinase complexes function in autophagy and carboxypeptidase Y sorting in *Saccharomyces cerevisiae*. *J. Cell Biol.* **152**, 519–530 (2001).
69. Stefan, C. J., Audhya, A. & Emr, S. D. The yeast synaptojanin-like proteins control the cellular distribution of phosphatidylinositol (4,5)-bisphosphate. *Mol. Biol. Cell* **13**, 542–557 (2002).
70. Roy, A. & Levine, T. P. Multiple pools of phosphatidylinositol 4-phosphate detected using the pleckstrin homology domain of Osh2p. *J. Biol. Chem.* **279**, 44683–44689 (2004).
71. Faulhammer, F. *et al.* Cell growth-dependent coordination of lipid signaling and glycosylation is mediated by interactions between Sac1p and Dpm1p. *J. Cell Biol.* **168**, 185–191 (2005).
72. Ivetac, I. *et al.* The type Ia inositol polyphosphate 4-phosphatase generates and terminates phosphoinositide 3-kinase signals on endosomes and the plasma membrane. *Mol. Biol. Cell* **16**, 2218–2233 (2005).
73. Munro, S. The Golgi apparatus: defining the identity of Golgi membranes. *Curr. Opin. Cell Biol.* **17**, 395–401 (2005).
74. Murray, D. & Honig, B. Electrostatic control of the membrane targeting of C2 domains. *Mol. Cell* **9**, 145–154 (2002).
75. Natarajan, P., Wang, J., Hua, Z. & Graham, T. R. Drs2p-coupled aminophospholipid translocase activity in yeast Golgi membranes and relationship to *in vivo* function. *Proc. Natl Acad. Sci. USA* **101**, 10614–10619 (2004).
76. Pomorski, T. *et al.* Drs2p-related P-type ATPases Dnf1p and Dnf2p are required for phospholipid translocation across the yeast plasma membrane and serve a role in endocytosis. *Mol. Biol. Cell* **14**, 1240–1254 (2003).
77. Ichimura, Y. *et al.* A ubiquitin-like system mediates protein lipidation. *Nature* **408**, 488–492 (2000).
78. Bijlmakers, M. J. & Marsh, M. The on-off story of protein palmitoylation. *Trends Cell Biol.* **13**, 32–42 (2003).
79. Smotryś, J. E. & Linder, M. E. Palmitoylation of intracellular signaling proteins: regulation and function. *Annu. Rev. Biochem.* **73**, 559–587 (2004).
80. Drenan, R. M. *et al.* Palmitoylation regulates plasma membrane-nuclear shuttling of R7BP, a novel membrane anchor for the RGS7 family. *J. Cell Biol.* **169**, 623–633 (2005).
81. Rocks, O. *et al.* An acylation cycle regulates localization and activity of palmitoylated Ras isoforms. *Science* **307**, 1746–1752 (2005).
82. Salcedo, S. P. & Holden, D. W. Bacterial interactions with the eukaryotic secretory pathway. *Curr. Opin. Microbiol.* **8**, 92–98 (2005).
83. Pizarro-Cerda, J. & Cossart, P. Subversion of phosphoinositide metabolism by intracellular bacterial pathogens. *Nature Cell Biol.* **6**, 1026–1033 (2004).
84. Meresse, S. *et al.* Controlling the maturation of pathogen-containing vacuoles: a matter of life and death. *Nature Cell Biol.* **1**, E183–E188 (1999).
85. Buttner, D. & Bonas, U. Port of entry — the type III secretion translocon. *Trends Microbiol.* **10**, 186–192 (2002).
86. Zhou, D. & Galan, J. *Salmonella* entry into host cells: the work in concert of type III secreted effector proteins. *Microbes Infect.* **3**, 1293–1298 (2001).
87. Gouin, E., Welch, M. D. & Cossart, P. Actin-based motility of intracellular pathogens. *Curr. Opin. Microbiol.* **8**, 35–45 (2005).
88. Via, L. E. *et al.* Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7. *J. Biol. Chem.* **272**, 13326–13331 (1997).
89. Vergne, I. *et al.* Mechanism of phagolysosome biogenesis block by viable *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **102**, 4033–4038 (2005).
90. Hernandez, L. D., Hueffer, K., Wenk, M. R. & Galan, J. E. *Salmonella* modulates vesicular traffic by altering phosphoinositide metabolism. *Science* **304**, 1805–1807 (2004).
91. Roy, C. R. Bacterial subversion of the host secretory pathway. *Proc. Natl Acad. Sci. USA* **102**, 1271–1272 (2005).
92. Nagai, H., Kagan, J. C., Zhu, X., Kahn, R. A. & Roy, C. R. A bacterial guanine nucleotide exchange factor activates ARF on *Legionella* phagosomes. *Science* **295**, 679–682 (2002).
93. Grieshaber, S. S., Grieshaber, N. A. & Hackstadt, T. *Chlamydia trachomatis* uses host cell dynein to traffic to the microtubule-organizing center in a p50 dynamitin-independent process. *J. Cell Sci.* **116**, 3793–3802 (2003).
94. Guignot, J. *et al.* Microtubule motors control membrane dynamics of *Salmonella*-containing vacuoles. *J. Cell Sci.* **117**, 1033–1045 (2004).
95. Harrison, R. E. *et al.* *Salmonella* impairs RILP recruitment to Rab7 during maturation of invasion vacuoles. *Mol. Biol. Cell* **15**, 3146–3154 (2004).
96. Boucrot, E., Henry, T., Borg, J. P., Gorvel, J. P. & Meresse, S. The intracellular fate of *Salmonella* depends on the recruitment of kinesin. *Science* **308**, 1174–1178 (2005).
97. Smith, G. A. & Enquist, L. W. Break ins and break outs: viral interactions with the cytoskeleton of mammalian cells. *Annu. Rev. Cell Dev. Biol.* **18**, 135–161 (2002).
98. Belov, G. A., Fogg, M. H. & Ehrenfeld, E. Poliovirus proteins induce membrane association of GTPase ADP-ribosylation factor. *J. Virol.* **79**, 7207–7216 (2005).
99. Loewen, C. J., Roy, A. & Levine, T. P. A conserved ER targeting motif in three families of lipid binding proteins and in Opi1p binds VAP. *EMBO J.* **22**, 2025–2035 (2003).
100. Sollner, T. *et al.* SNAP receptors implicated in vesicle targeting and fusion. *Nature* **362**, 318–324 (1993).
101. Banfield, D. K. SNARE complexes — is there sufficient complexity for vesicle targeting specificity? *Trends Biochem. Sci.* **26**, 67–68 (2001).
102. Short, B., Haas, A. & Barr, F. A. Golgins and GTPases, giving identity and structure to the Golgi apparatus. *Biochim. Biophys. Acta* **1744**, 383–395 (2005).

Acknowledgements We are indebted to M. Freeman, A. Gillingham, P. Langridge and K. Röper for comments on the manuscript. R.B. is supported by the Cambridge European Trust.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to S.M. (sean@mrc-lmb.cam.ac.uk).

Plasma membrane phosphoinositide organization by protein electrostatics

Stuart McLaughlin¹ & Diana Murray²

Phosphatidylinositol 4,5-bisphosphate (PIP₂), which comprises only about 1% of the phospholipids in the cytoplasmic leaflet of the plasma membrane, is the source of three second messengers, activates many ion channels and enzymes, is involved in both endocytosis and exocytosis, anchors proteins to the membrane through several structured domains and has other roles. How can a single lipid in a fluid bilayer regulate so many distinct physiological processes? Spatial organization might be the key to this. Recent studies suggest that membrane proteins concentrate PIP₂ and, in response to local increases in intracellular calcium concentration, release it to interact with other biologically important molecules.

If multiplicity of function is a fair measure of biological importance, phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂ or PIP₂) is the most important lipid in the cytoplasmic leaflet of the plasma membrane, its primary location in eukaryotic cells. This minor lipid (it constitutes about 1% of the lipids in the plasma membrane) is the source of three second messengers — inositol 1,4,5-trisphosphate (Ins(1,4,5)P₃), diacylglycerol (DAG) and PtdIns(3,4,5)P₃, and is responsible for a wide range of membrane-related phenomena (Fig. 1). How can one lipid regulate so many processes, apparently with spatial and temporal resolution? One answer leads to another question: why does the cell produce high concentrations of several 'natively unfolded' proteins and what are their functions? Biophysics provides some interesting clues that might help cell biologists answer both questions. Specifically, several natively unfolded proteins contain regions of basic/hydrophobic residues that bind PIP₂ electrostatically, then release it when the local intracellular Ca²⁺ level increases.

This review focuses on how electrostatic interactions can drastically affect the lateral distribution of PIP₂ in the membrane and how Ca²⁺/calmodulin (Ca/CaM) can modulate this distribution. We briefly review the importance of PIP₂, then consider theoretical and experimental studies showing that unstructured clusters of basic/hydrophobic residues on numerous membrane proteins (for example, ion channels, receptors and cytoskeletal proteins) produce a local positive electrostatic potential that acts as a strong basin of attraction for multivalent PIP₂ (valence $z = -4$ at pH 7), concentrating it in the 'electrical double layer' of 1 nm thickness adjacent to a basic cluster in the membrane. We then explore how, for many proteins, Ca/CaM binds to the cluster and removes it from the membrane, releasing the sequestered PIP₂. We conclude with some speculations about the biological importance of this reversible PIP₂ sequestration.

The enigma of PIP₂ functions

Figure 1 enumerates the many functions of PIP₂. For example, it is the source of the second messengers Ins(1,4,5)P₃, which releases Ca²⁺ from internal stores, and DAG, which acts together with Ca²⁺ to activate protein kinase C (PKC)^{1,2}. Well-characterized phosphoinositide-specific phospholipase Cs (PLCs) hydrolyse PIP₂ into these second mes-

sengers^{3,4}. For mammals, life begins with PIP₂ hydrolysis: a fertilizing spermatozoan injects not only its genetic material but also PLC- ζ into the egg, which produces a wave of Ins(1,4,5)P₃ and Ca²⁺ oscillations that raise the fertilization membrane and initiate egg activation, thereby launching embryo development⁵. In adult mammalian cells, phosphorylation of PIP₂ produces PtdIns(3,4,5)P₃, which acts as a second messenger by serving as a membrane anchor for several proteins⁶.

PIP₂ itself anchors proteins to the plasma membrane through pleckstrin homology (PH) and other domains of known structure^{7–10}. It activates at least two dozen different ion channels in the plasma membrane^{11,12}. It is required for clathrin-mediated endocytosis (refs 13–15; see also the review by McMahon & Gallop in this issue, p. 590) and is involved in phagocytosis^{15,16}, as well as for several stages of exocytosis^{14,17,18} and synaptic vesicle trafficking¹⁹. Furthermore, many proteins that regulate actin also bind PIP₂ (refs 20, 21), and laser-tweezer experiments show that PIP₂ functions as a second messenger to

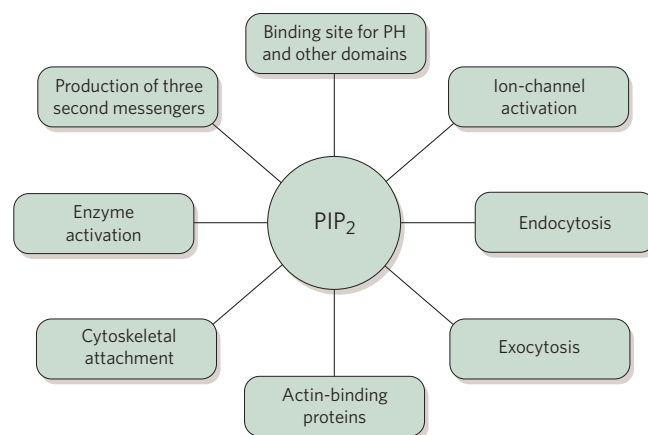


Figure 1 | Functions of PIP₂. See text for a brief discussion, or recent reviews for an extensive consideration, of the ion channels¹¹ and other proteins^{14,20} that interact with PIP₂. PH, pleckstrin homology.

¹Department of Physiology and Biophysics, Health Sciences Center, Stony Brook University, Stony Brook, New York 11794, USA; ²Department of Microbiology and Immunology, Weill Medical College of Cornell University, New York, New York 10021, USA.

regulate cytoskeleton–plasma-membrane adhesion²². Finally (but not exhaustively), it acts as a cofactor in enzyme activation²³. So how does PIP₂ do so much?

Irvine and colleagues suggest that there might be “multiple independent pools of PIP₂, with different pools governing distinct functions”²⁴. Emr, Stenmark and colleagues note that there might be “localized formation and turnover of this lipid”¹⁵. Recent reviews discuss the evidence for both localized production and sequestration of PIP₂ (ref. 20), and the controversial claim that PIP₂ is concentrated in lipid rafts^{25,26}. We focus on the biophysical evidence that suggests that the local level of free Ca²⁺ in a cell controls the local level of free PIP₂. Specifically, several proteins can sequester PIP₂ electrostatically, then release it in response to a local increase in Ca²⁺. This provides a mechanism that, if confirmed in cells, could validate the suggestion that there are multiple functionally independent pools of PIP₂ in the plasma membrane.

The cell has exquisitely tuned mechanisms to control the local free concentration of Ca²⁺ in the cytoplasm: “An extensive Ca²⁺-signalling toolkit is used to assemble signalling systems with very different spatial and temporal dynamics”²⁷. Our main thesis is that the cell uses this toolkit to help control the local level of free PIP₂ in particular regions of the plasma membrane, such as the growth cones of axons, the ruffles of fibroblasts and the nascent phagosomes of macrophages.

Cell biological, as well as biophysical, studies provide evidence that the natively unstructured proteins myristoylated alanine-rich C-kinase substrate (MARCKS) and growth-associated protein 43 (GAP43) are intimately involved with PIP₂, and act as ‘pipmodulins’ to control the level of free PIP₂ in the plasma membrane²⁸. Recent experiments provide important multifaceted evidence that MARCKS controls the level of free PIP₂ in the peripheral regions of cells (J. Sable and M. Sheetz, personal communication). For example, in cells treated with phorbol myristate acetate (PMA) to induce translocation of MARCKS from the plasma membrane to the cytoplasm, there is concomitant movement of a fluorescently labelled PLC- δ 1 PH domain, which binds specifically to PIP₂, from the cytoplasm to the plasma membrane. This strongly suggests that translocation of MARCKS produces an increase in the level of free PIP₂. Although many researchers — particularly structural biologists — often tacitly assume that function requires structure, more than 100 ‘natively unfolded’ or ‘intrinsically unstructured’ proteins have been identified; their evolutionary persistence strongly implies that they have an important function^{29,30}. MARCKS is a prototypical natively unfolded protein with a high fraction of negatively charged residues and few hydrophobic residues; thus, it exists in an extended conformation.

MARCKS acts as a reversible PIP₂ source

MARCKS is the major PKC substrate in most mammalian cells and is present at concentrations (1–10 μ M) comparable to those of PIP₂. Its properties have been reviewed elsewhere^{31–35}. Although it is highly negatively charged, MARCKS binds to the negatively charged cytoplasmic leaflet of the plasma membrane using two membrane anchors (Fig. 2a): an amino-terminal myristate inserts hydrophobically into the bilayer, and a conserved basic effector domain located in the middle of the molecule sticks electrostatically to the membrane³³. Neither the native MARCKS protein nor a peptide corresponding to the basic cluster requires PIP₂ for membrane binding: they bind strongly to membranes containing physiological fractions (15–30%) of monovalent acidic lipids, such as phosphatidylserine (PS)^{36–38}. When the membrane also contains PIP₂, both the native protein and the peptide laterally sequester the phosphoinositide. Fluorescence resonance energy transfer (FRET), electron paramagnetic resonance (EPR), self-quenching and PLC enzyme-activity measurements indicate that the 13 basic residues in the cluster laterally sequester three tetraivalent acidic PIP₂ by means of nonspecific electrostatics³⁹ (Fig. 2a).

Ca/CaM can bind to the unstructured basic cluster with high affinity (Fig. 2b). Stop-flow data imply that it interacts directly with the membrane-bound basic cluster and increases its rate of desorption⁴⁰. The crystal structure shows that Ca/CaM wraps around the central

region of the basic cluster (Fig. 3b), reversing its charge from positive to negative and repelling the complex from the membrane. The myristate anchor alone cannot hold the protein on the membrane, so Ca/CaM binding produces MARCKS translocation from the plasma membrane to the cytoplasm, and from lipid vesicles to solution³⁷, releasing the three electrostatically sequestered PIP₂ molecules (Fig. 2c).

MARCKS has a second mechanism for reversing PIP₂ sequestration. PKC-catalysed phosphorylation of three serine residues in the basic effector domain (Table 1) reduces its electrostatic attraction for the membrane, producing translocation of MARCKS from the plasma membrane to the cytoplasm in cells⁴¹ and from phospholipid vesicles to solution in a test tube^{36,37}. A corollary of both Ca/CaM-mediated and PKC-mediated MARCKS translocation from membrane to cytosol is the release of the three sequestered PIP₂s.

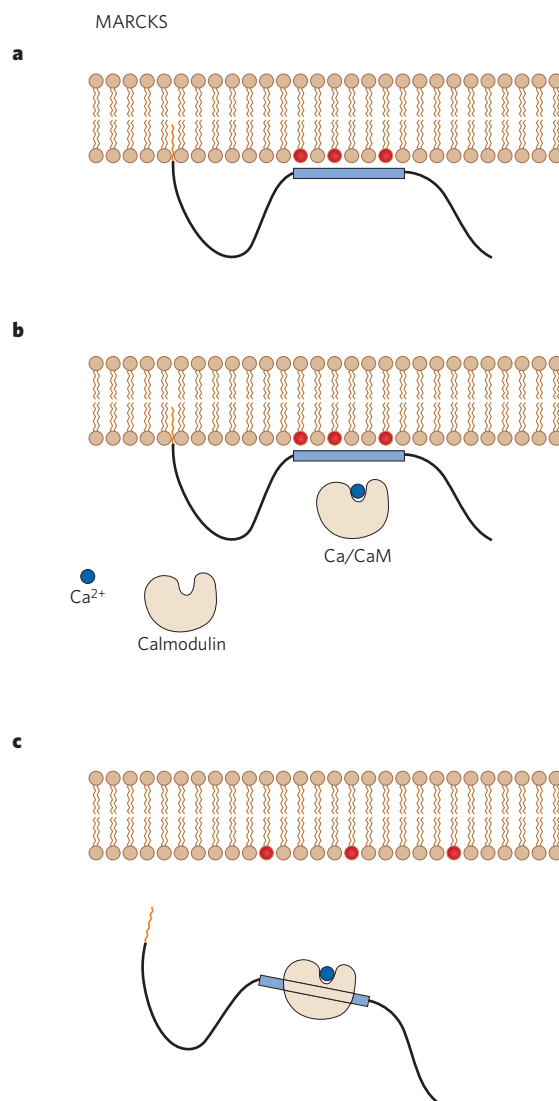


Figure 2 | MARCKS acts as a reversible source of PIP₂ in the plasma membrane. **a**, MARCKS attaches to the plasma membrane of a quiescent cell (low free cytoplasmic intracellular calcium concentration ([Ca²⁺]) and low protein kinase C (PKC) activity) by two anchors, an N-terminal myristate (orange) and a basic ‘effector domain’ (blue), which sequesters three PIP₂ molecules (lipids with red head groups). **b**, When the free cytoplasmic [Ca²⁺] increases, Ca/CaM binds to the basic cluster with high affinity and pulls it off the membrane, releasing three PIP₂ molecules. **c**, MARCKS bound to Ca/CaM translocates from the membrane to the cytoplasm. PKC phosphorylation of three serine residues within the basic effector domain also reduces its electrostatic interaction with the membrane and produces translocation of MARCKS from the plasma membrane to the cytoplasm.

Structural mechanisms of MARCKS, membranes and CaM

Figure 3a shows an atomic model of the basic/hydrophobic region of MARCKS overlaid on a bilayer. This represents the consistent picture of the protein–membrane interactions emerging from recent structural studies. EPR⁴² and circular dichroism (CD)⁴³ measurements show that the peptide is elongated both in solution and when bound to a membrane. EPR^{42,44} and nuclear magnetic resonance (NMR)^{45,46} experiments show that the five phenylalanine residues penetrate into

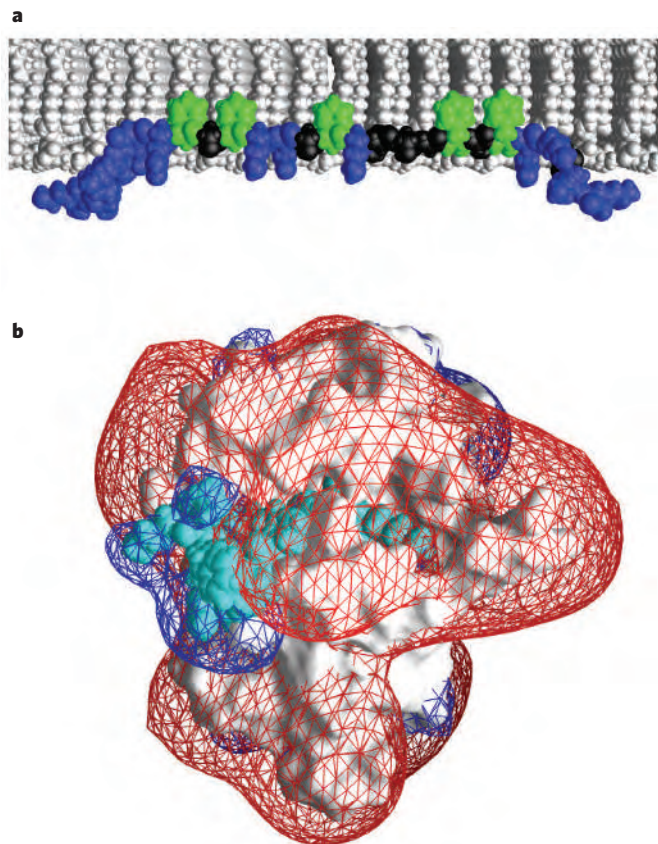


Figure 3 | The effector domain of MARCKS is mainly extended when bound to either a membrane or calmodulin. **a**, Atomic model of the basic effector domain (residues 151–175 of bovine MARCKS (Table 2) overlaid with a model of a lipid bilayer illustrating its experimentally determined extended conformation and the location of the five hydrophobic (green) and 13 basic (blue) residues. **b**, Crystal structure of the complex of a truncated effector domain (Table 1) with Ca/CaM. Reproduced with permission from Yamauchi *et al.*⁵⁰ (Protein Data Bank identifier: 1IWQ). The electrostatic equipotential profiles are shown in red (–25 mV) and blue (+25 mV) meshes. Note that the net negative charge of the complex will repel it from the negatively charged cytoplasmic leaflet of the plasma membrane (Fig. 5c). The potentials in Figs 3 and 5 were calculated using a modified version of Delphi⁷⁹ and visualized with Graphical Representation and Analysis of Structural Properties (GRASP)⁸⁰. [Salt] = 100 mM.

the hydrocarbon core⁴⁷ or acyl-chain region of the bilayer. Figure 3b illustrates the crystal structure of a complex of Ca/CaM with a slightly truncated peptide corresponding to the basic/hydrophobic region of MARCKS. Unlike complexes of other target peptides with Ca/CaM⁴⁸, the MARCKS basic region is mainly non-helical when bound to Ca/CaM, as shown by CD, NMR⁴⁹ and crystallography⁵⁰. Therefore, even the basic effector domain of MARCKS remains unfolded or extended in solution, when bound to the membrane and when bound to Ca/CaM. Nevertheless, studies in model systems and significant new data from living cells (M. Sheetz, personal communication) indicate that this region can perform the important function of sequestering and releasing PIP₂ in response to either Ca/CaM binding or PKC phosphorylation.

GAP43 might act as a reversible PIP₂ sink

Figure 4 illustrates how a different naturally unstructured protein, GAP43/neuromodulin^{51,52}, might act as a reversible ‘sink’ for PIP₂. The anchoring of GAP43 to the plasma membrane has been controversial⁵¹, but recent mass-spectrometric analysis⁵³ supports the earlier suggestion that at least a fraction of membrane-bound GAP43 contains two saturated acyl chains (for example, palmitate) attached to N-terminal cysteine residues⁵⁴. At physiological ionic strength, GAP43 binds both CaM and Ca/CaM with a modest dissociation constant (K_d) of ~1 μ M through its basic IQ (Table 1) domain. Liu and Storm⁵² proposed that GAP43, which is highly concentrated in the axonal growth cones of neurons, binds CaM in quiescent cells (low Ca²⁺ and low PKC activity); this binding holds the basic IQ domain off the membrane (Fig. 4a). They suggested that PKC phosphorylation of Ser41 in the IQ domain, which inhibits binding of CaM, releases high concentrations of CaM into the growth cone.

Recent important studies on the state of Ca/CaM in living cells^{55–57} imply that a similar release of CaM from GAP43 occurs when the level of Ca²⁺ increases, even without PKC phosphorylation of Ser41. These studies^{55–57} have provided convincing evidence that, in a typical quiescent cell, the level of free CaM (~10 μ M) is similar to the total level of CaM. When the cytoplasmic Ca²⁺ increases and most of the CaM is converted to Ca/CaM, however, the level of free Ca/CaM in a cell (~100 nM) is 100-fold lower than the total concentration. The intracellular concentration of proteins with a strong affinity for Ca/CaM is apparently sufficiently high to bind most of the Ca/CaM rapidly⁵⁵. Therefore, GAP43, which has weak (K_d = ~1 μ M) affinity for both CaM and Ca/CaM, should release CaM when the Ca²⁺ level rises. The corollary is that the released basic IQ domain should then bind to the membrane and laterally sequester PIP₂ (Fig. 4b). In other words, when the Ca²⁺ level in the growth cone of axons rises, or when PKC phosphorylates Ser41, GAP43 might act as a reversible PIP₂ sink (peptides corresponding to the GAP43 IQ domain do bind to PIP₂-containing vesicles⁵⁸).

Electrostatic interactions bind PIP₂ to basic clusters

Most lipids in the cytoplasmic leaflet of a mammalian plasma membrane are electrically neutral (for example, cholesterol), have zero net charge (for example, phosphatidylcholine or PC), or have a single net

Table 1 | The basic/hydrophobic regions discussed in this review and their approximate affinity for Ca²⁺/calmodulin

Region and sequence	K_d for Ca/CaM	Reference
MARCKS (154–171) KKRFS*FKKS*FKLSGFS*FK	10 nM	50
MARCKS (151–175) KKKKKRFS*FKKS*FKLSGFS*FKNNKK	10 pM†	Unpublished§
GAP43 (30–56) KAHKAATKIQAS*FRGHITRKLLKGEKK	1 μ M‡	52
NMDA NR1 (838–863) KRHKDARRKQMLAFAAVNVWRKNLQ	100 nM	67
NMDA NR1 (875–898) KKKATFRAITSLAS*S*FKRRRS*S*K	10 pM†	Unpublished§
EGFR (645–660) RRRHIVRKRT*LRLLQ	10 nM	65

Basic (R, K) residues are in blue and acidic (D, E) residues are in red. Protein kinase C phosphorylation of serine or threonine residues (designated by *) within the region decreases its affinity for Ca²⁺/calmodulin (CaM). †Values deduced from competition measurements of peptide binding to vesicles⁶⁵. For the bovine sequence myristoylated alanine-rich C-kinase substrate (MARCKS; 151–175) peptide, titration with dansyl-CaM indicated a dissociation constant (K_d) of 3 nM rather than 10 pM (ref. 84); peptide competition measurements confirmed that the K_d was about 10 nM for MARCKS (154–171) measured with dansyl-CaM⁶⁷. For the N-methyl-D-aspartate (NMDA) NR1 (875–898) peptide, titration with dansyl-CaM⁶⁷ indicated that the K_d was 4 nM rather than the 10 pM value deduced from competition measurements. The requirement for indirect competition measurements with these high affinity Ca/CaM complexes is discussed elsewhere⁴⁸. ‡Source: U. Golebiewska and S. McLaughlin, unpublished data. §CaM and Ca/CaM bind to GAP43 with similar affinity⁵², and the K_d for CaM determined in competition experiments with the GAP43 (30–56) peptide has a similar value. EGFR, epidermal growth factor receptor.

negative charge (for example, PS)⁵⁹. By contrast, polyphosphoinositides are highly charged. PIP₂ has a valence of -4 at pH 7.0 (ref. 60). Therefore, any protein with a cluster of four or more basic residues located at the membrane–solution interface should laterally sequester PIP₂ — that is, trap it in the electrical double-layer of 1 nm thickness in a manner similar to that illustrated in Fig. 5b. No structure is required for this lateral sequestration.

Figure 5a shows the predicted electrostatic equipotential surfaces (-25 mV; red mesh) adjacent to a 99:1 PC/PIP₂ bilayer. The equipotential profiles around each PIP₂ are approximately hemispheres; the potential is twice the value predicted by the Debye–Hückel theory because of the image-charge effect⁶¹. Therefore, the proximity of the low dielectric bilayer enhances electrostatic interactions at the membrane–solution interface. Figure 5b shows the adsorption of a peptide corresponding to the MARCKS basic effector domain to a 99:1 PC/PIP₂ membrane. Several different types of experiment have shown that the peptide sequesters three PIP₂ molecules electrostatically⁵⁸. The interaction is due to nonspecific electrostatics; for example, Lys 13 and Arg 13 peptides bind with identical affinity and the interaction is screened by salt. The $+13$ -valent basic peptides bind with about 100-fold stronger affinity to PC/PIP₂ vesicles than does the PH domain of PLC- δ 1, which forms a 1:1 ‘lock and key’ complex with PIP₂ with a K_d of about 1 μ M (refs 9, 58). Table 2 compares the binding characteris-

tics of the basic peptides with those of the PH domain.

Figure 5c shows the equipotential -25 -mV profile (red mesh) adjacent to a 2:1 PC/PS bilayer, which contains a similar fraction of monovalent acidic lipid to that of the cytoplasmic leaflet of a plasma membrane. The -25 -mV profile, located about 1 nm from the surface, is essentially flat. This application of electrostatic theory to an atomic model of the bilayer, where each atom has the appropriate partial charge, mirrors the predictions of the Gouy–Chapman theory⁶¹, which assumes that the charge is smeared uniformly over a structureless planar surface. This theory, and therefore the atomic model, is surprisingly consistent with a wide range of experiments on membranes⁶¹. Moreover, the atomic model adequately describes the electrostatic interaction of basic peptides with membranes^{62,63}. Figure 5d shows the -25 -mV (red) and $+25$ -mV (blue) electrostatic equipotential surfaces adjacent to a 2:1 PC/PS bilayer with an adsorbed MARCKS effector domain peptide: the calculations show that the basic residues produce a local positive potential (blue) that acts as a basin of attraction for multivalent PIP₂.

Figure 5d illustrates the essence of the electrostatic mechanism by which any membrane-bound cluster of basic residues can laterally sequester PIP₂. We treat PIP₂ as a -4 -valent point charge that does not perturb the potential; it therefore experiences an approximate potential of -25 mV when far from the cluster, but an approximate potential of $+25$ mV when adjacent to the cluster. The Boltzmann relation predicts that the local positive potential enhances the PIP₂ concentration about 1,000-fold. PIP₂ is not a point charge and does perturb the potential, but more realistic calculations using atomic models make a similar prediction: PIP₂ should be concentrated strongly (sequestered) adjacent to the basic cluster, even when the membrane has a 100-fold excess of monovalent acidic lipids (for example, 30% PS versus 0.3% PIP₂)⁶⁴. FRET and other experiments confirm this prediction for peptides corresponding to basic regions on MARCKS³⁹, epidermal growth factor receptor (EGFR)⁶⁵ and other proteins, such as the *N*-methyl-D-aspartate (NMDA) receptor.

Several other membrane proteins — for example, EGFR, NMDA receptor and the scaffolding protein Gravin/SSECKs — have apparently unfolded MARCKS-like basic/hydrophobic clusters that might bind to membranes and sequester PIP₂; Ca/CaM binding to these clusters should reverse the PIP₂ sequestration. Whereas certain regions of the cell have comparable concentrations of MARCKS or GAP43 and PIP₂, these other membrane proteins are present at lower concentrations. Therefore, their postulated reversible PIP₂ sequestration must function locally for the mechanism to be significant biologically. We consider how this might work for the two integral membrane proteins below.

NMDA and EGF receptors bind Ca/CaM

Figure 6a shows an illustration of the NMDA receptor⁶⁶: the NR1 subunit has two basic/hydrophobic regions in series (blue rectangles) that bind Ca/CaM⁶⁷. When glutamate binds to the channel (middle panel), Ca²⁺ enters the cell and Ca/CaM binds to the two clusters (right-hand diagram), contributing to inactivation of the channel by an unknown mechanism⁶⁷. These basic/hydrophobic clusters might also bind to the membrane and sequester PIP₂, as shown here. A peptide corresponding to the second high-Ca/CaM-affinity cluster does indeed bind to membranes containing PIP₂ (ref. 58), and Ca/CaM can displace the peptide from these membranes by binding to it with high affinity, $K_d \approx 10$ pM (Table 1).

Figure 6b shows an illustration of the EGFR. Structural studies have uncovered the mechanism by which binding of EGF to the exterior of the receptor produces dimerization⁶⁸, which leads to transautophosphorylation. Ligand binding produces a rapid increase in the free Ca²⁺ level⁶⁹. Model peptide studies using NMR and FRET indicate that the transmembrane helix breaks at the interface, and the basic/hydrophobic juxtamembrane region (eight basic residues; Table 1) is bound to the membrane, as shown in the illustration⁶⁵. Moreover, peptides corresponding to the juxtamembrane basic/hydrophobic cluster laterally

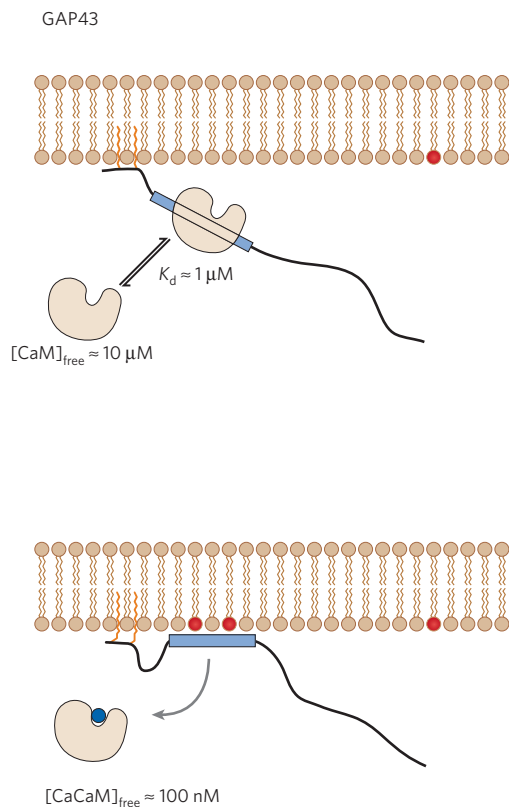


Figure 4 | GAP43 might act as a reversible PIP₂ sink in the axonal growth cones of neurons, where it is highly concentrated. **a**, Quiescent cell. Two palmitates (orange) anchor GAP43 to the plasma membrane. The free intracellular calmodulin concentration ($[CaM]_{free}$) is about 10 μ M, and the basic IQ domain (blue) binds CaM (and Ca/CaM) with a dissociation constant (K_d) of about 1 μ M (ref. 52). **b**, Activated cell with an elevated free cytoplasmic intracellular calcium concentration ($[Ca^{2+}]_{free}$): proteins with high-affinity binding sites bind about 99% of Ca/CaM. Consequently, raising $[Ca^{2+}]$ will cause CaM to desorb from the IQ domain, and GAP43 can act as a reversible source of CaM³². The basic IQ domain then binds to the adjacent negatively charged cytoplasmic leaflet and laterally sequesters PIP₂ (red lipids). $[Ca/CaM]_{free}$, free intracellular Ca^{2+}/CaM concentration.

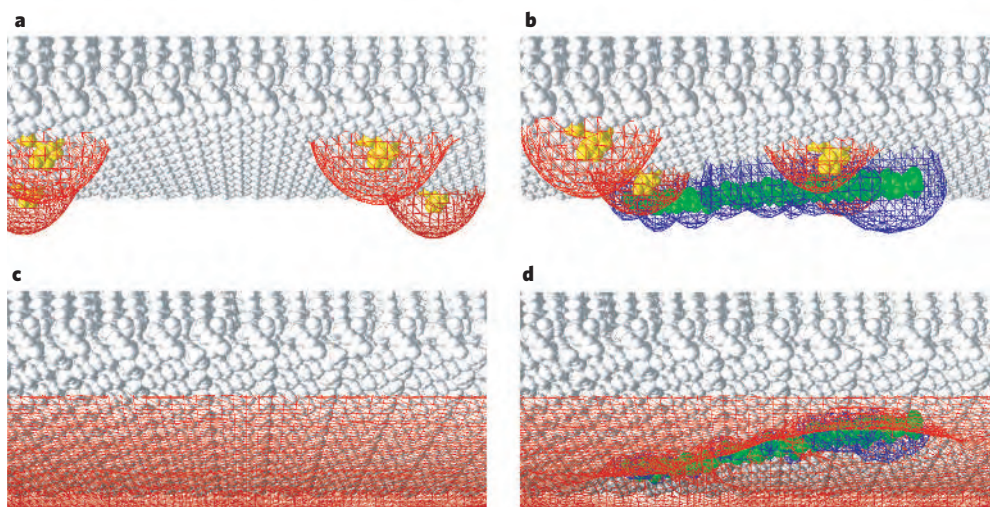


Figure 5 | The electrostatic mechanism by which membrane-bound clusters of basic residues laterally sequester the multivalent ($z = -4$) acidic lipid PIP_2 . **a**, Atomic model of a 99:1 phosphatidylcholine (PC)/ PIP_2 bilayer showing -25 mV electrostatic potentials (red mesh) and PIP_2 (yellow). **b**, On a 99:1 PC/ PIP_2 membrane, a peptide corresponding to MARCKS effector domain (residues 151–175; Table 2) forms a nonspecific electrostatic binding site for three PIP_2 molecules. **c**, Electrostatic potential adjacent to a 2:1 PC/PS membrane. The -25 mV equipotential profile (red) is located about 1 nm from the surface. **d**, On a 2:1 PC/PS membrane a MARCKS (151–175) peptide produces a local positive electrostatic potential (blue $= +25$ mV) that acts as a basin of attraction for PIP_2 , even when it is present at a (physiological) 100-fold lower mole fraction than PS.

sequester PIP_2 (left panel) and bind strongly to Ca/CaM, which releases the PIP_2 (right panel)⁶⁵. Phosphorylation of Thr 654 in the basic cluster inhibits Ca/CaM binding. The carboxy-terminal tail region of autophosphorylated EGFR binds two enzymes that use PIP_2 as a substrate: PLC- γ 1 and PtdIns-3-OH kinase PI(3)K. Therefore, EGFR might function as a scaffolding protein, binding PIP_2 and, on activation, releasing it to enzymes in the molecular neighbourhood that use PIP_2 as a substrate⁶⁵.

Can the behaviour of these basic/hydrophobic peptides be extrapolated to native proteins in the living cell? The available evidence strongly indicates that it can for MARCKS: if the protein translocates to the cytoplasm, it must release any sequestered PIP_2 , and membrane binding of the basic effector domain is required for anchorage, so it should sequester PIP_2 in the plasma membrane. The reversible sequestration mechanism is merely hypothetical for GAP43, NMDA receptor, EGFR and the many other proteins that have unstructured basic regions that might function in a similar manner.

Biological implications and future directions

Perhaps the most important recent advance in the phosphoinositide area has been the discovery that each intracellular organelle in a mammalian cell has a different array of kinases and phosphatases that produce and break down the phosphoinositides⁷⁰. This leads, for example, to the Golgi having a high level of PtdIns(4)P, but little PIP_2 . The phosphoinositides characteristic of each organelle (Fig. 4 in the review by Behnia & Munro, p. 596 in this issue) are intimately involved in membrane trafficking.

Despite its importance, we know little about the physical state of PIP_2 in the plasma membrane (or the other phosphoinositides in the

different intracellular organelles). Specifically, we do not know what fractions of PIP_2 are free and bound, where it is bound and how rapidly it diffuses. Several cell-biology studies provide evidence, based on the lateral distribution of the PLC- δ 1 PH domain probe, that PIP_2 is concentrated in ruffles^{71–73} and nascent phagosomes¹⁶; however, these reports are not universally accepted⁷⁴. MARCKS is concentrated in both ruffles and nascent phagosomes, possibly through its interaction with actin, but so are the kinases that produce PIP_2 (ref. 75). Therefore the relative importance of sequestration and synthesis remains unresolved. The potentially complicated biological interactions between Ca/CaM, PKC and basic clusters on MARCKS, GAP43, neurogranin, the NMDA NR1 subunit, the Ca^{2+} pump and so on, are discussed in an interesting essay⁷⁶. The relationship between putative cholesterol-enriched ‘rafts’ on the cytoplasmic leaflet and PIP_2 remains contentious^{25,26}. The role of gradients of PtdIns(3,4,5) P_3 in directing chemoattractant-induced pseudopod formation was recently reviewed⁷⁷; in this case, the gradients are probably caused by the non-uniform distributions of an important kinase (PI(3)K; concentrated at the leading edge) and phosphatase (PTEN; concentrated at the rear) that produce and break down PtdIns(3,4,5) P_3 .

We need new probes and approaches to determine the state of PIP_2 in the plasma membranes of living cells. For example, fluorescent PIP_2 , which works well for FRET and other experiments in model systems, is difficult to use in biological membranes because PIP_2 hydrolysis/turnover is rapid in a viable cell. Fluorescent non-hydrolysable PIP_2 analogues could make accurate lateral distribution and diffusion measurements trivial with existing technology. These do not need to be close chemical analogues, at least for testing whether electrostatic interactions produce a non-uniform distribution of the lipid in the

Table 2 | PIP_2 binding: PLC- δ 1 PH domain versus MARCKS basic effector domain

PLC- δ 1 PH domain	MARCKS basic cluster (residues 151–175)
Structured	Unstructured
Specific binding due to 12 hydrogen bonds	Non-specific electrostatic binding
Binds one PIP_2	Binds three PIP_2 molecules
K_d about 1 μM	Affinity for PC/ PIP_2 vesicles about 100-fold greater than PH domain
Function: anchors PLC to plasma membrane	Function: laterally sequesters PIP_2

K_d , dissociation constant; MARCKS, myristoylated alanine-rich C-kinase substrate; PH, pleckstrin homology; PIP_2 , phosphatidylinositol 4,5-bisphosphate; PLC, phospholipase C.

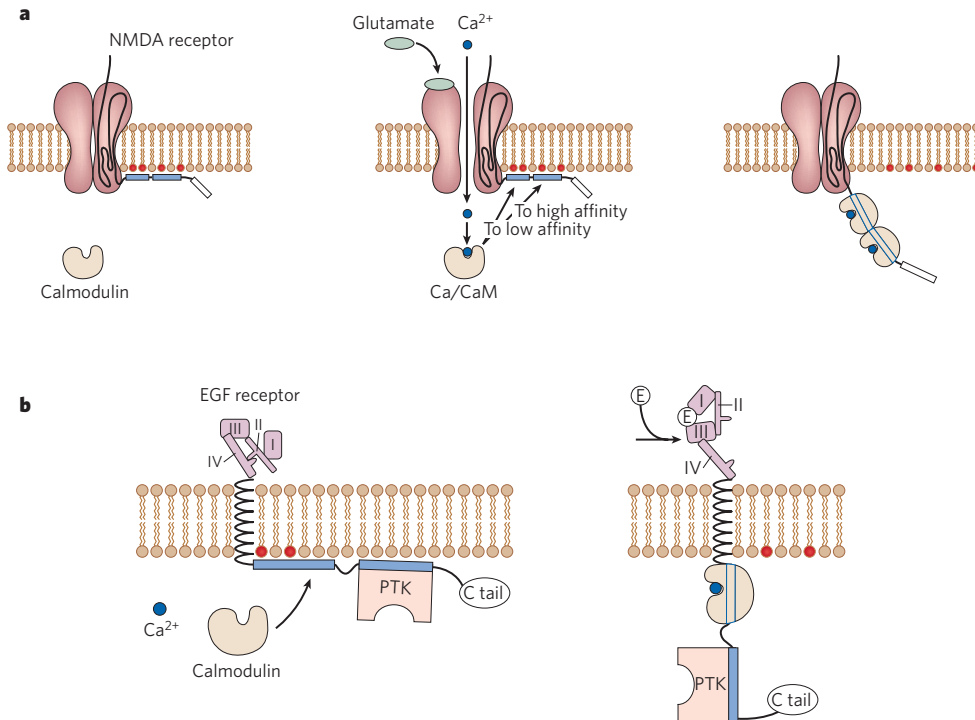


Figure 6 | Postulated reversible PIP_2 sequestration by the NMDA and epidermal growth factor (EGF) receptors. **a**, NMDA receptor. Two basic domains (blue) are located in series on the NR1 subunit; we postulate that they bind to membranes and laterally sequester PIP_2 (red lipids) by the same mechanism as myristoylated alanine-rich C-kinase substrate (MARCKS). Left: quiescent cell. Middle: glutamate binds to and activates the channel, allowing Ca^{2+} to flow into the cytoplasm where it forms Ca^{2+} /calmodulin (Ca/CaM), which then diffuses to the membrane-bound basic domains. Right: Ca/CaM binds with high affinity (about 10 pM) to the second (C1) basic domain, leveraging it from the membrane and allowing Ca/CaM to bind to the juxtamembrane domain, which has weaker affinity (about 100 nM). Binding of Ca/CaM helps inactivate the channel by an unknown mechanism. It also should release several laterally sequestered PIP_2 molecules. Figure adapted with permission from Ehlers *et al.*⁶⁷. **b**, EGF receptor (EGFR). The EGFR basic juxtamembrane (JM) region (blue rectangle) is postulated to bind reversibly to the plasma membrane^{65,81,82}. In a quiescent cell (left), the basic region binds to the plasma membrane and laterally sequesters PIP_2 (red) by nonspecific electrostatics; theoretical analysis also indicates that a basic face of the protein tyrosine kinase (PTK) domain (blue) binds to the negatively charged inner leaflet⁶⁵. Structural studies show that when EGF (E) binds to the extracellular domain, the dimerization arm on domain II is exposed⁶⁸. The JM, PTK domain and Ca/CaM are drawn to scale with the bilayer of thickness about 5 nm; the representation of the extracellular region, adapted with permission from Ferguson⁶⁸, is not to scale. Single-molecule studies (right) show that binding of EGF to only about 1% of the EGF receptors produces a robust increase in free intracellular calcium concentration⁶⁹. The JM region of EGFR binds Ca/CaM⁸³, which can rapidly pull the JM region from the membrane, releasing the sequestered PIP_2 (ref. 65).

plasma membrane: any fluorescent lipid with a valence $z = -4$ should faithfully reproduce the electrostatic interaction of PIP_2 with unstructured basic clusters. Comparison of the diffusion behaviour of the PIP_2 analogues with that of a more common and less highly charged lipid should also provide information about the state of PIP_2 in different regions of the membrane.

The PLC- $\delta 1$ PH domain has proved to be a valuable probe of PIP_2 ; however, it might also interact with membrane proteins when bound to PIP_2 and therefore report on both protein–probe and PIP_2 –probe interactions. New PH domain probes — that is, PH domains from other proteins — with high specificity for PIP_2 would be useful for comparative studies. For example, homology models imply that the PLC- $\delta 3$ PH domain should have a different charge density/chemical composition on its outer surface⁷⁸, and so should interact with membrane-bound proteins differently from the PLC- $\delta 1$ PH domain; this might be a useful high-affinity PIP_2 probe if it can be calibrated successfully in simple systems. With new fluorescent PIP_2 analogues, probes and approaches, such as single-molecule studies on living cells, we can begin to make real progress in understanding how PIP_2 exerts its many physiological functions. ■

1. Berridge, M. J. & Irvine, R. F. Inositol trisphosphate, a novel second messenger in cellular signal transduction. *Nature* **312**, 315–321 (1984).
2. Berridge, M. J. & Irvine, R. F. Inositol phosphates and cell signalling. *Nature* **341**, 197–205 (1989).
3. Rhee, S. G. Regulation of phosphoinositide-specific phospholipase C. *Annu. Rev. Biochem.*

70, 281–312 (2001).

4. Rebecchi, M. J. & Pentyala, S. N. Structure, function, and control of phosphoinositide-specific phospholipase C. *Physiol. Rev.* **80**, 1291–1335 (2000).
5. Swann, K., Larman, M. G., Saunders, C. M. & Lai, F. A. The cytosolic sperm factor that triggers Ca^{2+} oscillations and egg activation in mammals is a novel phospholipase C: PLC γ . *Reproduction* **127**, 431–439 (2004).
6. Cantley, L. C. The phosphoinositide 3-kinase pathway. *Science* **296**, 1655–1657 (2002).
7. Cho, W. & Stahlman, R. V. Membrane-protein interactions in cell signaling and membrane trafficking. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 119–151 (2005).
8. DiNitto, J. P., Cronin, T. C. & Lambright, D. G. Membrane recognition and targeting by lipid-binding domains. *Sci. STKE* 2003, re16 (2003).
9. Lemmon, M. A. Phosphoinositide recognition domains. *Traffic* **4**, 201–213 (2003).
10. Hurley, J. H. & Meyer, T. Subcellular targeting by membrane lipids. *Curr. Opin. Cell Biol.* **13**, 146–152 (2001).
11. Suh, B. C. & Hille, B. Regulation of ion channels by phosphatidylinositol 4,5-bisphosphate. *Curr. Opin. Neurobiol.* **15**, 370–378 (2005).
12. Hilgemann, D. W. Biochemistry. Oily barbarians breach ion channel gates. *Science* **304**, 223–224 (2004).
13. Honing, S. *et al.* Phosphatidylinositol-(4,5)-bisphosphate regulates sorting signal recognition by the clathrin-associated adaptor complex AP2. *Mol. Cell* **18**, 519–531 (2005).
14. Wenk, M. R. & De Camilli, P. Protein–lipid interactions and phosphoinositide metabolism in membrane traffic: insights from vesicle recycling in nerve terminals. *Proc. Natl Acad. Sci. USA* **101**, 8262–8269 (2004).
15. Simonsen, A., Wurms, A. E., Emr, S. D. & Stenmark, H. The role of phosphoinositides in membrane transport. *Curr. Opin. Cell Biol.* **13**, 485–492 (2001).
16. Botelho, R. J. *et al.* Localized biphasic changes in phosphatidylinositol-4,5-bisphosphate at sites of phagocytosis. *J. Cell Biol.* **151**, 1353–1368 (2000).
17. Holz, R. W. *et al.* A pleckstrin homology domain specific for phosphatidylinositol 4,5-bisphosphate (PtdIns-4,5- P_2) and fused to green fluorescent protein identifies plasma membrane PtdIns-4,5- P_2 as being important in exocytosis. *J. Biol. Chem.* **275**, 17878–17885 (2000).
18. Martin, T. F. PI(4,5) P_2 regulation of surface membrane traffic. *Curr. Opin. Cell Biol.* **13**, 493–499 (2001).

19. Di Paolo, G. *et al.* Impaired PtdIns(4,5)P₂ synthesis in nerve terminals produces defects in synaptic vesicle trafficking. *Nature* **431**, 415–422 (2004).
20. Janmey, P. A. & Lindberg, U. Cytoskeletal regulation: rich in lipids. *Nature Rev. Mol. Cell Biol.* **5**, 658–666 (2004).
21. Papayannopoulos, V. *et al.* A polybasic motif allows N-WASP to act as a sensor of PIP₂ density. *Mol. Cell* **17**, 181–191 (2005).
22. Raucher, D. *et al.* Phosphatidylinositol 4,5-bisphosphate functions as a second messenger that regulates cytoskeleton-plasma membrane adhesion. *Cell* **100**, 221–228 (2000).
23. Sciorra, V. A. *et al.* Dual role for phosphoinositides in regulation of yeast and mammalian phospholipase D enzymes. *J. Cell Biol.* **159**, 1039–1049 (2002).
24. Hincliffe, K. A., Ciruela, A. & Irvine, R. F. PIPkins, their substrates and their products: new functions for old enzymes. *Biochim. Biophys. Acta* **1436**, 87–104 (1998).
25. van Rhee, J., Achame, E. M., Janssen, H., Calafat, J. & Jalink, K. PIP₂ signaling in lipid domains: a critical re-evaluation. *EMBO J.* **24**, 1664–1673 (2005).
26. Golub, T. & Caroni, P. PI(4,5)P₂-dependent microdomain assemblies capture microtubules to promote and control leading edge motility. *J. Cell Biol.* **169**, 151–165 (2005).
27. Berridge, M. J., Bootman, M. D. & Roderick, H. L. Calcium signalling: dynamics, homeostasis and remodelling. *Nature Rev. Mol. Cell Biol.* **4**, 517–529 (2003).
28. Laux, T. *et al.* GAP43, MARCKS, and CAP23 modulate PI(4,5)P₂ at plasmalemmal rafts, and regulate cell cortex actin dynamics through a common mechanism. *J. Cell Biol.* **149**, 1455–1472 (2000).
29. Dunker, A. K. *et al.* Intrinsically disordered protein. *J. Mol. Graph. Model.* **19**, 26–59 (2001).
30. Uversky, V. N. What does it mean to be natively unfolded? *Eur. J. Biochem.* **269**, 2–12 (2002).
31. Aderem, A. The MARCKS brothers: a family of protein kinase C substrates. *Cell* **71**, 713–716 (1992).
32. Blackshear, P. J. The MARCKS family of cellular protein kinase C substrates. *J. Biol. Chem.* **268**, 1501–1504 (1993).
33. McLaughlin, S. & Aderem, A. The myristoyl-electrostatic switch: a modulator of reversible protein-membrane interactions. *Trends Biochem. Sci.* **20**, 272–276 (1995).
34. Arbuza, A., Schmitz, A. A. & Vergeres, G. Cross-talk unfolded: MARCKS proteins. *Biochem. J.* **362**, 1–12 (2002).
35. Sundaram, M., Cook, H. W. & Byers, D. M. The MARCKS family of phospholipid binding proteins: regulation of phospholipase D and other cellular components. *Biochem. Cell Biol.* **82**, 191–200 (2004).
36. Taniguchi, H. & Manenti, S. Interaction of myristoylated alanine-rich protein kinase C substrate (MARCKS) with membrane phospholipids. *J. Biol. Chem.* **268**, 9960–9963 (1993).
37. Kim, J., Shishido, T., Jiang, X., Aderem, A. & McLaughlin, S. Phosphorylation, high ionic strength, and calmodulin reverse the binding of MARCKS to phospholipid vesicles. *J. Biol. Chem.* **269**, 28214–28219 (1994).
38. Rusu, L., Gambhir, A., McLaughlin, S. & Radler, J. Fluorescence correlation spectroscopy studies of peptide and protein binding to phospholipid vesicles. *Biophys. J.* **87**, 1044–1053 (2004).
39. Gambhir, A. *et al.* Electrostatic sequestration of PIP₂ on phospholipid membranes by basic/aromatic regions of proteins. *Biophys. J.* **86**, 2188–2207 (2004).
40. Arbuza, A., Murray, D. & McLaughlin, S. MARCKS, membranes, and calmodulin: kinetics of their interaction. *Biochim. Biophys. Acta* **1376**, 369–379 (1998).
41. Ohmori, S. *et al.* Importance of protein kinase C targeting for the phosphorylation of its substrate, myristoylated alanine-rich C-kinase substrate. *J. Biol. Chem.* **275**, 26449–26457 (2000).
42. Qin, Z. & Cafiso, D. S. Membrane structure of protein kinase C and calmodulin binding domain of myristoylated alanine rich C kinase substrate determined by site-directed spin labeling. *Biochemistry* **35**, 2917–2925 (1996).
43. Wang, J., Arbuza, A., Hangyas-Mihalyne, G. & McLaughlin, S. The effector domain of myristoylated alanine-rich C kinase substrate binds strongly to phosphatidylinositol 4,5-bisphosphate. *J. Biol. Chem.* **276**, 5012–5019 (2001).
44. Victor, K., Jacob, J. & Cafiso, D. S. Interactions controlling the membrane binding of basic protein domains: phenylalanine and the attachment of the myristoylated alanine-rich C-kinase substrate protein to interfaces. *Biochemistry* **38**, 12527–12536 (1999).
45. Zhang, W., Crocker, E., McLaughlin, S. & Smith, S. O. Binding of peptides with basic and aromatic residues to bilayer membranes: phenylalanine in the myristoylated alanine-rich C kinase substrate effector domain penetrates into the hydrophobic core of the bilayer. *J. Biol. Chem.* **278**, 21459–21466 (2003).
46. Ellena, J. F., Burnitz, M. C. & Cafiso, D. S. Location of the myristoylated alanine-rich C-kinase substrate (MARCKS) effector domain in negatively charged phospholipid bilicles. *Biophys. J.* **85**, 2442–2448 (2003).
47. White, S. H. & Wimley, W. C. Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 319–365 (1999).
48. Clapperton, J. A., Martin, S. R., Smerdon, S. J., Gamblin, S. J. & Bayley, P. M. Structure of the complex of calmodulin with the target sequence of calmodulin-dependent protein kinase I: studies of the kinase activation mechanism. *Biochemistry* **41**, 14669–14679 (2002).
49. Porumb, T., Crivici, A., Blackshear, P. J. & Ikura, M. Calcium binding and conformational properties of calmodulin complexed with peptides derived from myristoylated alanine-rich C kinase substrate (MARCKS) and MARCKS-related protein (MRP). *Eur. Biophys. J.* **25**, 239–247 (1997).
50. Yamauchi, E., Nakatsu, T., Matsubara, M., Kato, H. & Taniguchi, H. Crystal structure of a MARCKS peptide containing the calmodulin-binding domain in complex with Ca²⁺-calmodulin. *Nature Struct. Biol.* **10**, 226–231 (2003).
51. Benowitz, L. I. & Routtenberg, A. GAP-43: an intrinsic determinant of neuronal development and plasticity. *Trends Neurosci.* **20**, 84–91 (1997).
52. Liu, Y. C. & Storm, D. R. Regulation of free calmodulin levels by neuromodulin: neuron growth and regeneration. *Trends Pharmacol. Sci.* **11**, 107–111 (1990).
53. Liang, X., Lu, Y., Neubert, T. A. & Resh, M. D. Mass spectrometric analysis of GAP-43/neuromodulin reveals the presence of a variety of fatty acylated species. *J. Biol. Chem.* **277**, 33032–33040 (2002).
54. Skene, J. H. & Virag, I. Posttranslational membrane attachment and dynamic fatty acylation of a neuronal growth cone protein, GAP-43. *J. Cell Biol.* **108**, 613–624 (1989).
55. Persechini, A. & Stemmer, P. M. Calmodulin is a limiting factor in the cell. *Trends Cardiovasc. Med.* **12**, 3–37 (2002).
56. Black, D. J., Tran, Q. K. & Persechini, A. Monitoring the total available calmodulin concentration in intact cells over the physiological range in free Ca²⁺. *Cell Calcium* **35**, 415–425 (2004).
57. Isotani, E. *et al.* Real-time evaluation of myosin light chain kinase activation in smooth muscle tissues from a transgenic calmodulin-biosensor mouse. *Proc. Natl Acad. Sci. USA* **101**, 6279–6284 (2004).
58. Wang, J. *et al.* Lateral sequestration of phosphatidylinositol 4,5-bisphosphate by the basic effector domain of myristoylated alanine-rich C kinase substrate is due to nonspecific electrostatic interactions. *J. Biol. Chem.* **277**, 34401–34412 (2002).
59. Holthuis, J. C. & Levine, T. P. Lipid traffic: floppy drives and a superhighway. *Nature Rev. Mol. Cell Biol.* **6**, 209–220 (2005).
60. McLaughlin, S., Wang, J., Gambhir, A. & Murray, D. PIP₂ and proteins: interactions, organization, and information flow. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 151–175 (2002).
61. McLaughlin, S. The electrostatic properties of membranes. *Annu. Rev. Biophys. Biomol. Struct.* **18**, 113–136 (1989).
62. Ben-Tal, N., Honig, B., Peitzsch, R. M., Denisov, G. & McLaughlin, S. Binding of small basic peptides to membranes containing acidic lipids: theoretical models and experimental results. *Biophys. J.* **71**, 561–575 (1996).
63. Murray, D., Arbuza, A., Honig, B. & McLaughlin, S. The role of electrostatic and nonpolar interactions in the association of peripheral proteins with membranes. *Curr. Top. Membr.* **52**, 271–302 (2002).
64. Wang, J., Gambhir, A., McLaughlin, S. & Murray, D. A computational model for the electrostatic sequestration of PI(4,5)P₂ by membrane-adsorbed basic peptides. *Biophys. J.* **86**, 1969–1986 (2004).
65. McLaughlin, S., Smith, S. O., Hayman, M. J. & Murray, D. An electrostatic engine model for autoinhibition and activation of the epidermal growth factor receptor (EGFR/Erbb) family. *J. Gen. Physiol.* **126**, 41–53 (2005).
66. Wollmuth, L. P. & Sobolevsky, A. I. Structure and gating of the glutamate receptor ion channel. *Trends Neurosci.* **27**, 321–328 (2004).
67. Ehlers, M. D., Zhang, S., Bernhardt, J. P. & Huganir, R. L. Inactivation of NMDA receptors by direct interaction of calmodulin with the NR1 subunit. *Cell* **84**, 745–755 (1996).
68. Ferguson, K. M. Active and inactive conformations of the epidermal growth factor receptor. *Biochem. Soc. Trans.* **32**, 742–745 (2004).
69. Uyemura, T., Takagi, H., Yanagida, T. & Sako, Y. Single-molecule analysis of epidermal growth factor signaling that leads to ultrasensitive calcium response. *Biophys. J.* **88**, 3720–3730 (2005).
70. De Matteis, M. A. & Godi, A. PI-loting membrane traffic. *Nature Cell Biol.* **6**, 487–492 (2004).
71. Honda, A. *et al.* Phosphatidylinositol 4-phosphate 5-kinase 4 is a downstream effector of the small G protein ARF6 in membrane ruffle formation. *Cell* **99**, 521–532 (1999).
72. Tall, E. G., Spector, I., Pentyala, S. N., Bitter, I. & Rebecchi, M. J. Dynamics of phosphatidylinositol 4,5-bisphosphate in actin-rich structures. *Curr. Biol.* **10**, 743–746 (2000).
73. Watt, S. A., Kular, G., Fleming, I. N., Downes, C. P. & Lucocq, J. M. Subcellular localization of phosphatidylinositol 4,5-bisphosphate using the pleckstrin homology domain of phospholipase C δ 1. *Biochem. J.* **363**, 657–666 (2002).
74. van Rhee, J. & Jalink, K. Agonist-induced PIP₂ hydrolysis inhibits cortical actin dynamics: regulation at a global but not at a micrometer scale. *Mol. Biol. Cell* **13**, 3257–3267 (2002).
75. Doughman, R. L., Firestone, A. J. & Anderson, R. A. Phosphatidylinositol phosphate kinases put PI4,5P₂ in its place. *J. Membr. Biol.* **194**, 77–89 (2003).
76. Chakravarthy, B., Morley, P. & Whitfield, J. Ca²⁺-calmodulin and protein kinase Cs: a hypothetical synthesis of their conflicting convergences on shared substrate domains. *Trends Neurosci.* **22**, 12–16 (1999).
77. Van Haastert, P. J. & Devreotes, P. N. Chemotaxis: signalling the way forward. *Nature Rev. Mol. Cell Biol.* **5**, 626–634 (2004).
78. Singh, S. M. & Murray, D. Molecular modeling of the membrane targeting of phospholipase C pleckstrin homology domains. *Protein Sci.* **12**, 1934–1953 (2003).
79. Gallagher, K. & Sharp, K. Electrostatic contributions to heat capacity changes of DNA-ligand binding. *Biophys. J.* **75**, 769–776 (1998).
80. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).
81. Hunter, T., Ling, N. & Cooper, J. A. Protein kinase C phosphorylation of the EGF receptor at a threonine residue close to the cytoplasmic face of the plasma membrane. *Nature* **311**, 480–483 (1984).
82. Ullrich, A. *et al.* Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature* **309**, 418–425 (1984).
83. Li, H., Ruano, M. J. & Villalobo, A. Endogenous calmodulin interacts with the epidermal growth factor receptor in living cells. *FEBS Lett.* **559**, 175–180 (2004).
84. Graff, J. M., Young, T. N., Johnson, J. D. & Blackshear, P. J. Phosphorylation-regulated calmodulin binding to a prominent cellular substrate for protein kinase C. *J. Biol. Chem.* **264**, 21818–21823 (1989).
85. Yarniuk, A. P. & Vogel, H. J. Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Mol. Biotechnol.* **27**, 33–57 (2004).

Acknowledgements This work was supported by grants from the NIH and Carol M. Baldwin Foundation (to S.M.) and NSF (to D.M.).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to S.M. (stuart.mclaughlin@stonybrook.edu).

Role of cholesterol and lipid organization in disease

Frederick R. Maxfield¹ & Ira Tabas²

Membrane lipids are essential for biological functions ranging from membrane trafficking to signal transduction. The composition of lipid membranes influences their organization and properties, so it is not surprising that disorders in lipid metabolism and transport have a role in human disease. Significant recent progress has enhanced our understanding of the molecular and cellular basis of lipid-associated disorders such as Tangier disease, Niemann–Pick disease type C and atherosclerosis. These insights have also led to improved understanding of normal physiology.

The lipid components of biological membranes are important for normal cell function, and their improper distribution or metabolism can have serious consequences for cells and organisms. Some of the important functions of membranes — such as providing a permeability barrier that separates compartments in eukaryotic cells — have been appreciated since the first observations of subcellular organelles. Other functions, such as signalling by phosphoinositides, have also been studied for decades, but recent advances indicate new ways in which these signalling mechanisms can be regulated both spatially and temporally. In the past few years, several lines of evidence have shown that the biophysical properties of membrane bilayers have significant effects on the properties of membrane proteins¹.

Changes in the organization of lipids can have profound effects on cellular functions such as signal transduction and membrane trafficking^{2–4}. These membrane effects can cause disease in humans as a result of genetic alterations or environmental effects (such as diet), or both. Cholesterol is one of the most important regulators of lipid organization, and mammals have developed sophisticated and complex mechanisms to maintain cellular cholesterol levels in membranes within a narrow range⁵. When these homeostatic mechanisms are overwhelmed, as in the late stages of atherosclerosis, the consequences can be severe.

Our understanding of the contributions of membranes to disease varies from disorder to disorder. The role of cholesterol and lipids in atherosclerosis has been studied for decades⁶, and many of the cellular and molecular mechanisms have been worked out in considerable detail. This has been one of the leading examples (perhaps the best example) of how modern tools of cell and molecular biology can result in understanding and treatment of human disease. However, even in this case, there are important unresolved questions about how cholesterol affects cells in atherosclerotic lesions, how cholesterol moves within cells and how cholesterol is exported to extracellular acceptors.

With other disorders, such as the inherited lysosomal storage diseases (which lead to lipid accumulation in cells), the molecular defects have been identified, but it is often not clear how these defects lead to the particular set of symptoms that afflict patients or how to relieve these symptoms. In other cases, there are tantalizing hints that membrane organization is important, but the details remain very uncertain.

For example, polymorphisms in the apolipoprotein Apo-E are strongly linked with the age of onset of Alzheimer's disease, but the basis for this linkage remains unclear. Similarly, treatment with cholesterol-lowering statins has been reported to have beneficial effects in delaying the average onset of Alzheimer's disease, but the cellular and molecular basis for these effects are not clear⁷.

In this review, we will briefly summarize the current state of knowledge of membrane organization and lipid trafficking in mammalian cells. We then discuss how changes in lipid composition and organization can lead to altered cell function, and, where possible, we will relate this to our understanding of the pathophysiology associated with these disorders.

Membrane organization

The membranes of mammalian cells have several functional roles that must be carried out simultaneously. The membranes provide a permeability barrier that allows different ion and solute concentrations to exist on each side of the membrane. This allows specialized functions in various organelles and maintains transmembrane electrical potentials. At the same time, membranes provide a scaffold for supporting membrane proteins and yet they must be fluid enough to allow rapid diffusion of these proteins. Membranes must also be flexible enough to bend, for example, when budding to form vesicles or tubules, or fusing with other membranes during trafficking. It is now understood that certain lipids, especially the phosphoinositides, are used by cells to organize signal-transduction processes at certain locations within cells. Furthermore, many biological membranes have a lateral inhomogeneity (microdomains) that can be used to help bring signalling molecules (both lipids and membrane proteins) together or to keep them apart under various conditions^{2–4}.

The competing demands of these functions place stringent restrictions on the lipid compositions of membranes. For example, highly ordered membranes can generally provide a better permeability barrier than more disordered membranes because polar molecules can more easily intercalate into the disordered lipids. However, highly ordered (gel-phase) lipids would not allow rapid diffusion of membrane proteins, and they would be difficult to bend into vesicles and tubules. In fact, gel-phase lipids are not observed in mammalian cell

¹Department of Biochemistry, Weill Medical College of Cornell University, 1300 York Avenue, New York, NY 10021, USA; ²Departments of Medicine, Cell Biology, and Physiology & Cellular Biophysics, Columbia University College of Physicians and Surgeons, New York, NY, USA.

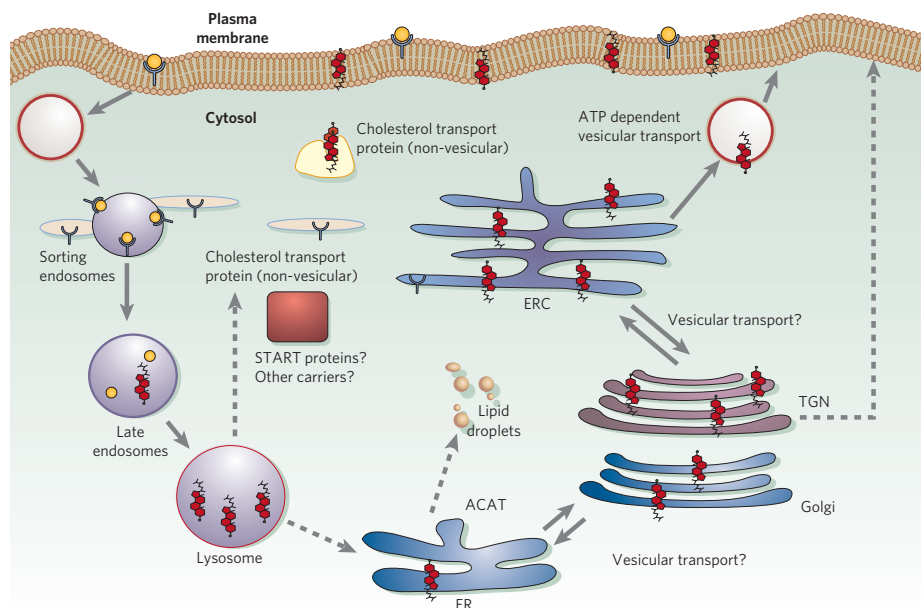


Figure 1 | Intracellular cholesterol transport. LDL (yellow circles) carrying cholesterol and cholesterol esters bound to LDL receptors (light blue Y-shape) is internalized and transported to sorting endosomes and to late endosomes and lysosomes from which cholesterol can efflux to cellular compartments including the plasma membrane or the endoplasmic reticulum (ER). The LDL receptor recycles to the plasma membrane via the endocytic recycling compartment (ERC). Efflux from late endosomes and lysosomes is poorly characterized as indicated by the dashed lines. Cholesterol can move from the plasma membrane to the ERC by a non-vesicular, ATP-independent process. Recycling of cholesterol back to the plasma membrane occurs by non-vesicular transport and in membrane-recycling vesicles carrying other recycling membrane components. Newly synthesized cholesterol in the ER is mostly transported from the ER directly to the plasma membrane, bypassing the Golgi, but some follows the biosynthetic secretory pathway from the ER to the Golgi. Excess cholesterol in the ER becomes esterified by ACAT and stored in cytoplasmic lipid droplets. TGN, trans-Golgi network.

membranes, which exhibit more dynamic liquid-phase properties. In biological membranes, there is often a mixture of liquid-disordered (l_d) and liquid-ordered (l_o) phases (see Box 1) with the abundance of these types dependent upon lipid composition.

Two or more lipid phases can coexist within a single bilayer, and the importance of this in biological membranes has been studied intensely in the past few years^{2,3,8}. The plasma membrane has been studied most, and there is considerable evidence that small domains (microdomains) coexist with different lipid organizations. At present, however, there is considerable uncertainty about the abundance, size, duration and composition differences of the different types of microdomain, and there remain some unanswered questions about whether they exist at all⁹. On the basis of lipid composition, sensitivity to detergent extraction and biophysical measurements of the motion of lipid analogues, it seems that a large fraction, and perhaps the majority, of the lipids in the plasma membrane are in an l_o type of organization³.

The different types of lipid organization have important effects on membrane proteins. First, many membrane proteins prefer to associate with a particular type of organization, and this will lead to their physical separation in bilayers with coexisting lipid phases^{8,10}. For proteins with lipid or fatty-acid anchors, this preference will be determined largely by the properties of the acyl chains. For example, glycosylphosphatidylinositol (GPI)-anchored proteins have a preference for ordered domains because their phosphoinositide anchors typically have saturated acyl chains¹¹. Similarly, many palmitoylated and myristoylated proteins (for example, Src-family kinases) associate with ordered domains in the cytoplasmic leaflet¹². Prenylated membrane anchors (for example, on Ras superfamily GTPases) have a preference for disordered domains because of the unsaturation of the isoprenyl groups^{13,14}. It should be pointed out, however, that these preferences on the basis of hydrophobic membrane anchors may be outweighed by interactions with other proteins or by headgroup interactions.

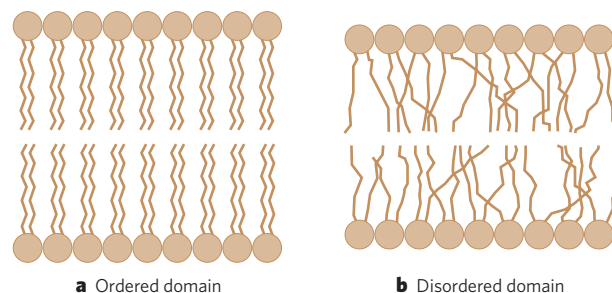
Transmembrane proteins also have preferences for l_o or l_d membranes^{8,10}. Most transmembrane proteins seem to have a preference for l_d domains, according to evidence of their sensitivity to mild detergent extraction. Physically this may be related to the difficulty of accommodating a protein transmembrane domain within a tightly packed lipid bilayer without disturbing the lipid organization. Nevertheless, some transmembrane proteins show a preference for more ordered domains, and this is often more pronounced when the proteins are

crosslinked⁴. This organizational preference is important in signal-transduction processes and in sorting during membrane trafficking. More ordered membranes have thicker lipid bilayers (see Fig. 1), and this can lead to preferential inclusion of membrane proteins with longer hydrophobic sequences in their transmembrane domain. Organizational preferences have been proposed to influence the sorting of proteins in the secretory pathway^{10,15}, although this view has recently been challenged by measurements of average bilayer thickness in isolated organelles¹⁶.

One of the most important proposed roles for membrane domains is in the regulation of signal transduction. In particular, crosslinking some signalling receptors can lead to formation of signalling complexes that are associated with ordered, raft-like membranes^{4,8}. One of the best-characterized examples is the crosslinking of IgE receptors, which are found on mast cells and are involved in triggering allergic reactions. Crosslinking of these receptors increases their resistance to solubilization by cold Triton X-100, indicating that they are recruited to more ordered membrane domains. Another signalling protein

Box 1 | Ordered and disordered lipids

In ordered lipid phases, the atoms of the acyl chains are tightly packed and relatively elongated (Box 1 Fig. 1a). The disordered lipid phases (l_d) are characterized by rapid diffusion in the plane of the membrane and a poorly ordered structure in the hydrophobic core of the bilayer (Box 1 Fig. 1b). Lipids with unsaturated fatty acids, which have kinks in their acyl chains, increase the propensity of a bilayer to be in an l_d organization. The atoms in this phase are not tightly packed, which allows water molecules and other small molecules to penetrate into the bilayer relatively easily. For similar reasons, this type of lipid organization is very susceptible to solubilization by mild detergents⁹². The thickness of the bilayer also decreases as the acyl chains of the lipids become disordered.



a Ordered domain

b Disordered domain

recruited to these ordered domains is the Src-family kinase Lyn. Recent studies indicate that Lyn in these ordered domains is protected from an inactivating transmembrane phosphatase, and this leads to greater net phosphorylation of the IgE receptor⁴. Depletion of cellular cholesterol prevents the recruitment of Lyn and the IgE receptor to detergent-resistant ordered domains and abrogates signalling. Placing the cytosolic domain of protein tyrosine phosphatase α on a myristoyl and palmitoyl anchor, which allows it to enter ordered membrane domains, prevented phosphorylation of crosslinked IgE receptors⁴. These results suggest that lipid microdomains are important for segregating proteins on the basis of their preferences for different types of lipid order. This is one way that signalling can be regulated.

In many other cases, the mechanisms by which lipid ordering affects signalling is less clear. Cholesterol is an important determinant of membrane organization (Box 2), and it is relatively easy to manipulate cellular cholesterol levels in cell culture. Thus, cholesterol depletion is often used as a method to test whether lipid organization plays

a part in a signalling pathway. A limitation of this is that the levels of cholesterol depletion are often non-physiological, and nonspecific effects can be caused by severe cholesterol depletion. Nevertheless, cholesterol depletion, along with other data such as detergent resistance, can indicate a role for lipid organization in signalling. Several studies have suggested that there is a relationship between lipid domains and the actin cytoskeleton⁴. Moderate reductions in cholesterol cause an inhibition of neutrophil motility and a marked reduction in actin-dependent protrusions on human neutrophils in response to chemoattractants, and similar results are seen in other cell types¹⁷. Some of these effects are associated with reduced activation of the small GTPase Rac in cells, with an approximately 30% reduction in cholesterol, but the precise mechanism linking lipid organization to Rac signalling is not known.

Membrane bilayer properties can affect the activity of single proteins — particularly proteins with multiple membrane-spanning domains that undergo conformational changes as part of their activity cycle¹. For example, the Ca^{2+} -ATPase from the sarcoplasmic reticulum, which pumps Ca^{2+} into the sarcoplasmic reticulum of muscle cells, can be reconstituted into vesicles made up of phospholipids with various fatty-acid lengths. The ATPase activity is greatest when the lipids approximately match the bilayer thickness of cellular membranes. The changes in bilayer thickness would be expected to affect the differences in free energy of the different conformational states of the protein as it goes through its activity cycle, and this could affect the kinetics of the pump activity. The activity of the $\text{Na}^{+}, \text{K}^{+}$ -ATPase reconstituted into artificial membranes can be altered by the addition of cholesterol to the membranes, but it is uncertain whether this is due to changes in biophysical properties of the bilayer or to specific associations with cholesterol¹. For several proteins, a specific sterol-sensing domain has been identified. In the sterol-regulatory-element-binding protein (SREBP) cleavage-activating protein (SCAP), there is a particular sequence in a transmembrane domain that is responsible for binding cholesterol and inducing a conformational change. This allows SCAP and its associated SREBP to be transported out of the endoplasmic reticulum (ER) when cellular cholesterol is low¹⁸. These results emphasize that lipids can have effects on protein function both through specific binding and through changes in the bilayer biophysical properties.

Different organelles within a cell have distinct lipid compositions. The plasma membrane has a high concentration of cholesterol, whereas the outer leaflet of the plasma membrane has high levels of sphingomyelin and glycosphingolipids. This composition is consistent with a highly ordered membrane. The endocytic recycling compartment¹⁹ and parts of the *trans*-Golgi⁸ are also relatively highly ordered membranes. At the opposite extreme, the ER has a low cholesterol content and a large fraction of unsaturated lipids, which contribute to a more disordered membrane organization. It is unclear how this type of lipid organization contributes to the function of the ER. One possibility is that it facilitates activities such as the flipping of dolichol-linked glycoconjugates and certain lipid precursors from the cytosolic leaflet to the luminal side of the ER membrane.

Cholesterol is synthesized in the ER and delivered to other organelles by a combination of vesicular and non-vesicular transport processes^{20–22} (Fig. 1). The mechanisms of non-vesicular cholesterol transport are only partly understood, but there is substantial evidence that this is the major route for cholesterol movement between organelles. Because cholesterol is very insoluble in water, it must be shuttled by carriers. A few candidates for cholesterol carriers in the cytoplasm have been identified²², but in most cases proof of their role is still lacking. Perhaps the best-documented example is the steroidogenic acute regulatory protein (StAR), which is the prototype for the StAR-related lipid transfer (START) family of transport proteins. StAR plays an essential role in the delivery of cholesterol to mitochondria, where it is used in steroid hormone synthesis in steroidogenic tissues. Other members of this family of proteins can bind cholesterol or other lipids and facilitate their intracellular, non-vesicular transport²³. Stud-

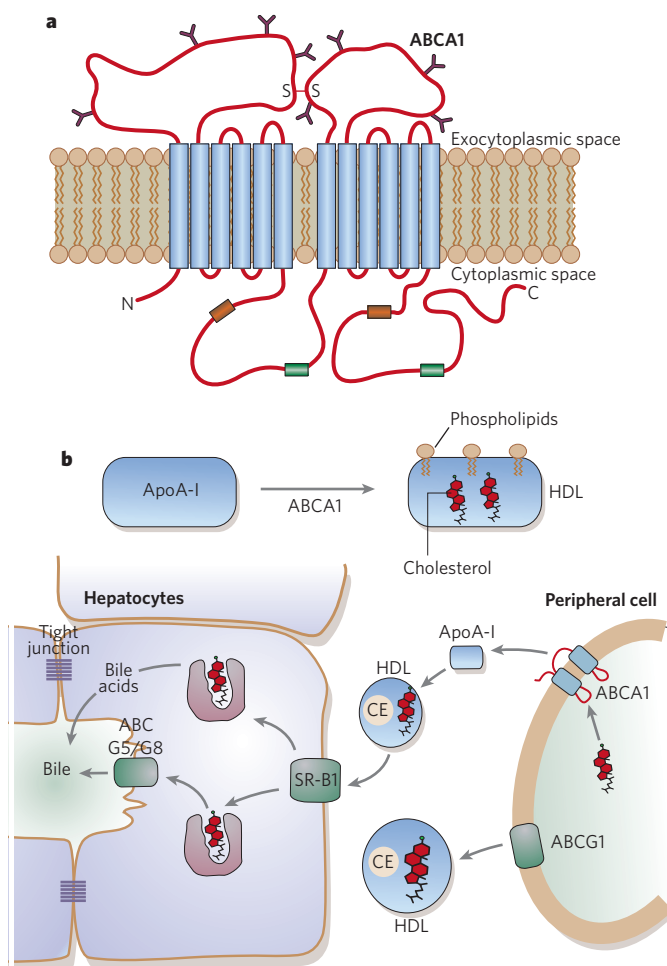


Figure 2 | Cholesterol efflux. **a**, ABCA1 is a multiple membrane-spanning protein with two nucleotide-binding folds linked by a cytoplasmic peptide sequence. **b**, ABCA1 promotes the transfer of phospholipids to lipid-poor forms of ApoA-I, the major protein component of HDLs. The mechanisms for this transfer are not fully understood, but ABCA1 apparently functions by translocating phospholipids and perhaps cholesterol across the plasma membrane bilayer and presenting them to ApoA-I, which binds to ABCA1 (refs 39, 43). **c**, These ApoA-I particles are transformed into HDLs in the blood by the action of lyssolecithin:cholesterol acyltransferase (LCAT). ABCG1 promotes the efflux of cholesterol to larger HDL particles^{99,100}. The HDLs bind to scavenger receptor-B1 (SR-B1) in hepatocytes, and transfer their associated cholesterol and cholesterol esters to the liver. The cholesterol is excreted into the bile either as free cholesterol (by way of ABCG5/G8) or after conversion to bile salts⁴³.

ies of the transport of a naturally fluorescent sterol, dehydroergosterol (DHE), have shown that sterols can have a very high flux through the cytoplasm^{19,24}. For example, in Chinese hamster ovary cells, in which about 40% of the DHE is in the endocytic recycling compartment, the DHE in that compartment can be replenished with a $t_{1/2}$ of 2–3 min-utes after photobleaching¹⁹.

The other major source of cellular cholesterol is endocytic uptake of lipoproteins such as low-density lipoprotein (LDL) and hydrolysis of their cholesterol esters in late endosomes and lysosomes⁵. The lipoprotein-derived cholesterol is rapidly released from these hydrolytic organelles and delivered throughout the cell. Studies of Niemann–Pick disease type C (NPC), an inherited lysosomal storage disorder that leads to accumulation of cholesterol and other lipids, have shown that a luminal protein (NPC2; ref. 25) and a transmembrane protein (NPC1; ref. 26) in late endosomes are required for efflux of cholesterol from these organelles, but the details of how these proteins work remain to be determined^{25,27,28}. In normal cells, these efflux mechanisms keep the cholesterol content in late endosome membranes low.

Homeostatic mechanisms

Organisms must maintain the proper functioning of their membranes in response to various changes. In humans, one of the most significant factors affecting the membrane is dietary intake of cholesterol and fats, which are delivered to cells throughout the body through lipoproteins⁵. Because cholesterol is a major regulator of lipid organization, its cellular concentration must be maintained within a narrow range, and cells have a variety of mechanisms for accomplishing this. Rapid transport of sterol by vesicular and non-vesicular pathways ensures that any changes are rapidly reflected in changes in cholesterol levels in many organelles.

A rapid response to increasing cholesterol levels is the esterification of excess cholesterol by an ER enzyme, acyl CoA:cholesterol acyl-transferase (ACAT)²⁹. The esterified cholesterol is stored in cytoplasmic lipid droplets. The cholesterol esters in the droplets are hydrolysed by neutral cholesterol ester hydrolases, which in some cells include a hormone-sensitive lipase that also hydrolyses triglycerides in fat cells³⁰. The cholesterol released from the droplets can be used for cell membranes and, in steroidogenic cells, for steroid hormone synthesis. The cycle of cholesterol esterification and hydrolysis may provide the major short-term buffering of cholesterol levels in cells. The activity of ACAT is regulated by cholesterol levels³¹, providing a homeostatic sensor. This regulation occurs at two levels: ACAT is allosterically regulated by cholesterol³², and cholesterol loading of cells promotes more rapid efflux of sterol from the plasma membrane²⁴, which could increase the rate of delivery to ACAT in the ER³³.

As mentioned earlier, many genes involved in cholesterol metabolism are regulated by SREBP³⁴. When cholesterol levels are high, SREBP and SCAP are retained in the ER by binding to INSIG, a resident ER protein. When cholesterol is low, the SREBP–SCAP complex exits from the ER, and SREBP undergoes two proteolytic cleavages. This releases the cytosolic domain of SREBP, which is then translocated into the nucleus and regulates the transcription of many genes, including the LDL receptor and HMG-CoA reductase, the rate-limiting enzyme in cholesterol synthesis. Thus, this system regulates both the synthesis of cholesterol and its uptake through lipoproteins.

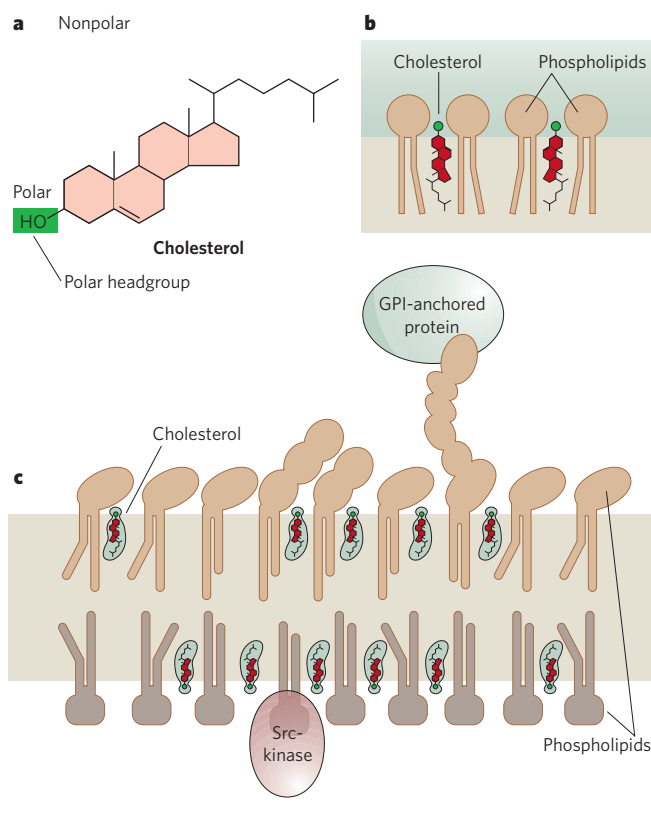
A third level of cholesterol regulation within cells is provided by cholesterol efflux mechanisms (Fig. 2). In the past few years, the molecular mechanisms for cellular export of cholesterol have begun to be understood, but there are still significant gaps in our knowledge concerning how these processes operate and how they are regulated. The key extracellular acceptors for cholesterol are high-density lipoproteins (HDLs) and one of their associated apolipoproteins, ApoA-I (ref. 35). The family of ABC transporters has a key role in delivering cholesterol and phospholipids to apolipoproteins, and defects in these transporters lead to several human diseases, as discussed below³⁶. Indeed, the identification of the gene responsible for Tangier disease, which

leads to HDL deficiency and increased cholesterol ester storage in macrophages, was the key to discovering the role of the ABC transporters in cholesterol efflux. The lipid-bilayer properties influence cholesterol efflux, but the precise mechanisms for this are unclear.

In several studies, cholesterol efflux has been linked to expression of caveolin, the coat protein of caveolae, and it has been proposed that these cholesterol-rich membrane domains may be a site of cholesterol

Box 2 | Effects of cholesterol on lipid organization.

The structure of cholesterol is very different from that of other membrane lipids (Box 2 Fig. 1a). The body of cholesterol consists of a series of fused rings, which make that part of the molecule quite rigid. At one end of this planar ring system is a hydroxyl group and at the other end is a hydrocarbon tail, so cholesterol, like other membrane lipids, has both hydrophilic and hydrophobic poles that determine its positioning within the lipid bilayer. When the hydroxyl group is next to the phospholipid ester carbonyl, the rigid body of cholesterol is situated alongside the fatty-acid tails of neighbouring phospholipids and can help to order these tails. Cholesterol can have preferential interactions with certain lipids, either because its small headgroup requires additional shielding from adjacent lipids⁹³ or through hydrogen-bonded complexes with lipids such as sphingomyelin⁹⁴ (Box 2 Fig. 1b). The polar moiety of cholesterol is much smaller than the polar headgroups of other lipids, so flip-flop between leaflets of the membrane bilayer can occur readily. The transbilayer distribution of cholesterol is not known. Cholesterol can increase the order in liquid membranes through the effects of its rigid ring system and the ability to fill interstitial spaces. A particularly important type of lipid organization is the l_o phase⁹⁵ in which the atoms in the hydrophobic core are more tightly packed than in the l_d phase, but the lipid molecules are able to diffuse in the plane of the bilayer almost as rapidly as in the l_d phase. The l_o organization provides a good permeability barrier while allowing movement of membrane constituents. Cholesterol is an important component of l_o phase membranes, and its structure seems to allow it to fill interstitial spaces between lipids (providing tight packing), while still allowing rapid diffusion. Cholesterol may also help to stabilize boundaries between coexisting lipid domains⁹⁶. In model membranes other lipids with small polar head groups (for example, ceramide) may be able to also support the formation of an l_o -type of organization^{97,98}. The l_o type of lipid order would be associated with detergent-resistant lipid domains or 'rafts'. The ordering of the acyl chains also causes thickening of the bilayer.



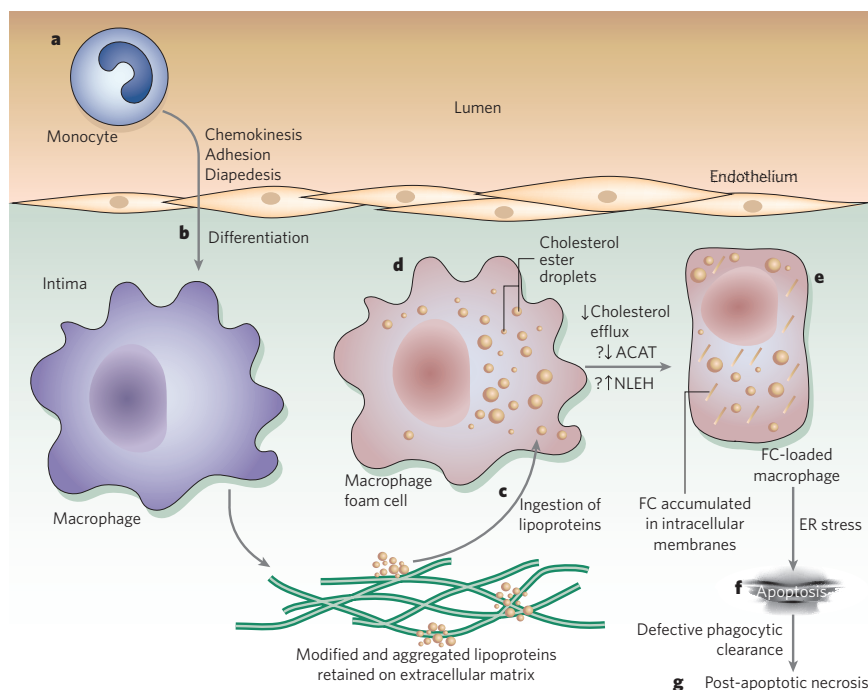


Figure 3 | Entry and cholesterol loading of macrophages in atherosclerotic lesions. **a**, Monocytes are attracted to focal areas of the arterial wall in which atherogenic lipoproteins have been retained on the extracellular matrix. These retained lipoproteins, particularly those whose phospholipids are modified by oxidation, signal to the endothelium to express chemokines and adhesion molecules. **b**, The monocytes then migrate through the endothelial layer and differentiate into macrophages. **c**, The macrophages ingest the retained lipoproteins by endocytic and phagocytic mechanisms and thus acquire a large load of lipoprotein-derived cholesterol. **d**, In early lesions, the cholesterol is stored as ACAT-derived cholesteryl esters and thus acquire a foamy appearance. **e**, In advanced lesions, unesterified or 'free' cholesterol (FC) accumulates, leading to macrophage apoptosis (**f**) and necrosis (**g**).

efflux³⁷. How this works is unclear. In cultured cells, overexpression of stearoyl CoA desaturase leads to an increase in unsaturated acyl chains in membrane lipids and a decrease in the resistance of the plasma-membrane lipids to extraction with cold Triton X-100 (refs 38, 39). These changes in bilayer properties are associated with a decrease in cholesterol efflux to apo A-I but an increase in passive release of cholesterol to an HDL2 acceptor^{38,39}.

In most peripheral tissues, metabolism of cholesterol (other than esterification/de-esterification) is a minor biochemical pathway. However, most cells convert a small fraction of cholesterol into oxysterols⁴⁰, and these molecules are important for intracellular signalling. For example, cholesterol overload in cells, which activates the nuclear liver X receptor/retinoid X receptors (LXR/RXR), perhaps through oxysterol intermediates, triggers a 'reverse cholesterol transport' programme involving both cellular cholesterol efflux and transport of the effluxed cholesterol to the liver for secretion in bile⁴¹. Specifically, activated LXR/RXR leads to the induction of the efflux receptors ABCA1 and ABCG1, the efflux enhancer apolipoprotein E, the plasma lipid transfer proteins CETP and PLTP, stearoyl CoA desaturase, the bile synthetic enzyme Cyp7a and the cholesterol-to-bile transporter ABCG5/G8 (refs 35, 42, 43). In the liver, cholesterol is excreted into the bile both as free cholesterol and after conversion to bile acids.

In addition to regulation of membrane properties by changes in cholesterol, the degree of unsaturation of the acyl chains in phospholipids is an important determinant of membrane biophysical properties. The availability of fatty acids for incorporation into phospholipids is controlled by several factors, including their extracellular availability (for example, through dietary sources), and by a complex network of metabolic regulatory mechanisms. The SREBP family of transcriptional regulators has significant effects on levels of proteins involved in fatty-acid synthesis and modification in addition to their effects on cholesterol synthesis⁴⁴. Furthermore, SREBP as well as other transcriptional regulators of fatty-acid metabolism such as the LXRs, the peroxisome proliferator-activated receptors (PPARs) and hepatocyte nuclear factor can all be regulated by polyunsaturated fatty acids^{44,45}. This regulation affects the amount of fatty acids in the cell and the degree of their unsaturation, which can alter the saturation of the fatty acids incorporated into phospholipids as well as the cholesterol:phospholipid ratio in the cell. This regulation can help to maintain the proper biophysical properties of the cell membranes, but it is uncer-

tain whether membrane bilayer properties themselves directly regulate these biosynthetic pathways.

Disorders of lipid and cholesterol metabolism

Although alterations in lipid metabolism and distribution may contribute to many diseases, there are several genetic diseases for which alterations in lipid traffic or metabolism are the primary cause. The study of these disorders has contributed enormously to our understanding of basic mechanisms of lipid metabolism and transport. At the same time, the detailed aetiology of these diseases is often difficult to explain.

Under normal conditions, membrane components that are delivered to late endosomes and lysosomes are subject to hydrolysis by the hydrolytic enzymes in these organelles. This catabolic process is important for the normal turnover of lipid components, and a lack of activity from one of these hydrolases leads to an accumulation of the undegraded substrate for the missing hydrolase. Several lysosomal storage disorders (including Tay-Sachs, Fabry, Niemann-Pick type A or B, and Sandhoff diseases) arise from defects in the breakdown of lipids in late endosomes and lysosomes. These can be caused by defects in a single hydrolytic enzyme or in activator proteins that participate in the digestion of sphingolipids⁴⁶. Normally, the sphingolipids and glycosphingolipids become segregated into internal membranes in late endosomes along with an unusual negatively charged lipid, bis(monoacylglycerol)phosphate (BMP), which is also called lysobisphosphatidic acid (LBPA)⁴⁷. These internal membranes can be continuous with the limiting membrane or detached to form internal vesicles. In either case, they appear as 'multi-vesicular bodies' in electron micrographs. When hydrolysis of sphingolipids is impaired, the internal membranes containing sphingolipids and BMP accumulate, and almost the entire lumen of the organelles can become filled with these membranes, which can form a series of internal membrane whorls⁴⁶.

NPC disease is a lysosomal storage disorder that shares many characteristics with the lysosomal sphingolipid enzyme deficiencies. It is caused by a primary defect in cholesterol or lipid trafficking rather than an enzymatic deficiency^{28,48}. Late endosomes lacking functional forms of either the NPC1 protein or the NPC2 protein show very slow efflux of cholesterol from late endosomes. Because NPC2 is a late endosome luminal protein that binds cholesterol⁴⁹, it is likely that cholesterol transport is the primary defect in cells carrying this mutation.

For the NPC1 protein, which is a multi-spanning membrane protein, the mechanism by which it affects cholesterol transport is not known. Internal membranes accumulate in the late endosomes of NPC cells. Additionally, there are changes in the trafficking of various lipid molecules and cholesterol that are similar in NPC and several of the lipid hydrolysis enzyme deficiencies⁵⁰. In effect, these storage organelles become a sink for lipids and cholesterol in the cell. Changes in lipid composition or cholesterol content can alter endocytic sorting of lipids³, but it is not known precisely how such changes exert their effects. It is plausible that a buildup of one membrane lipid component traps certain other lipids in the internal membrane whorls of the storage organelles. Interestingly, overexpression of Rab7 or Rab9, small GTPases that regulate vesicle trafficking, can partly correct the NPC phenotype of cholesterol storage in tissue culture fibroblasts⁵⁰.

In most lysosomal storage diseases, the accumulation of lipids can be seen in many tissues, and it can be observed in cultured fibroblast lines from the affected individuals. Typically, the most serious effects are seen in the brain, which leads eventually to neuronal death and neurological complications that are often severe and frequently fatal. In some cases, the defects appear at early developmental stages. It is unclear precisely how the various types of lysosomal storage disorder lead to cell death. Perhaps dysfunctional catabolism by lysosomes leads to shortages of certain metabolites. Another possibility is that blockage of vesicle transport along microtubules is blocked by the enlarged storage organelles; or maybe levels of important signalling molecules (for example, oxysterols) are reduced because their precursors are not released into the cytoplasm. It is unclear at present if any of these are the important causes of pathology or whether cell death arises by some other mechanism.

Defects in members of the ABC family of transporters are associated with a variety of human diseases. Tangier disease is a very rare autosomal recessive disorder caused by defects in ABCA1 (ref. 51). This is associated with a severe deficiency in HDL and reduced efflux of cholesterol, especially from macrophages and other reticuloendothelial cells. This leads to cholesterol ester accumulation in these cells and is also associated with increased susceptibility to atherosclerosis. ABCG5 and ABCG8 are expressed in the liver and intestines⁵¹. These proteins can transport cholesterol and other sterols into the bile or the intestines. In the intestines ABCG5 and ABCG8 excrete newly

absorbed plant sterols to a much greater extent than cholesterol. Defects in these proteins lead to a rare autosomal recessive disorder, sitosterolaemia, that is associated with a large increase in plasma levels of plant sterols but only modest increases in plasma cholesterol⁵¹. The buildup in plant sterols, such as sitosterol, is associated with tendon and tuberous xanthomas as well as arthritis and atherosclerosis.

Early events in atherosclerosis

Atherosclerosis is the major human disease associated with cholesterol and lipid metabolism. The earliest detectable event in atherogenesis (the process of forming atheromas) is the accumulation of plasma lipoproteins in the subendothelium, or intima, of focal areas of the arterial tree⁵². The lipoproteins are retained owing to a combination of proteoglycan binding and lipoprotein aggregation, which impedes egress from the arterial wall because of their increased particle size. These retained lipoproteins, particularly those that are modified by oxidation, aggregation and other means, elicit a series of biological responses that lead to atherogenesis⁵². Chief among these biological responses is an unusual type of inflammation consisting of infiltration of monocytes and T cells but not neutrophils⁵³ (Fig. 3). More specifically, certain types of oxidized phospholipid derived from modified lipoproteins can activate the overlying endothelium to secrete chemokines and express adhesion molecules for monocytes and T cells⁵⁴. These leukocytes migrate through an otherwise intact endothelial layer, and the monocytes eventually differentiate into macrophages in the intima⁵⁵.

Once embedded in the intima, the macrophages encounter native and modified lipoproteins, most of which are bound to the matrix. Through a process that is only partly understood but has aspects related to receptor-mediated endocytosis and phagocytosis, the macrophages ingest the lipoprotein particles^{56–59}. In a cell-culture model of the initial interaction of macrophages with retained and aggregated lipoproteins, significant rearrangement of the actin cytoskeleton and protrusion of membrane processes is seen, and this is required for the continued uptake of cholesterol into the cells⁵⁹. Interestingly, just as cholesterol depletion inhibits signal-dependent actin assembly in some cells¹⁷, loading of macrophages with cholesterol through uptake of modified lipoproteins or through a cyclodextrin carrier (that is, without a lipoprotein) can lead to increased actin assembly and protrusion of membrane processes in macrophages⁶⁰. It seems likely that these cholesterol-dependent effects are mediated by changes in lipid organization, which can affect activation of the small GTPase Rac^{17,60}. In the blood vessel wall, the initial contact with lipoproteins could lead to cholesterol transfer to the macrophages, leading to actin-dependent protrusions. This would enhance the further uptake of cholesterol into the cells.

Most of the cholesterol in lipoproteins is in the form of cholesteryl fatty-acyl esters. These esters are hydrolysed to cholesterol and fatty acids in acidic, degradative organelles such as late endosomes and then transported to other sites in the macrophage. Cholesterol transport to the plasma membrane is important for cholesterol efflux; transport to the ER is necessary for intracellular cholesterol homeostasis (through SREBP) and for re-esterification by ACAT. Transport to the mitochondria leads to the formation of oxysterols, which, in turn, may have roles in LXR activation and sterol efflux. ACAT-mediated re-esterification is a major fate of lipoprotein-derived cholesterol in intimal macrophages. The resulting cholesteryl ester molecules coalesce into membrane-bound neutral lipid droplets in the cytoplasm, a feature that has given rise to the term 'foam cell'⁶¹.

Receptor-mediated uptake by means of the LDL receptor is usually limited because of its homeostatic downregulation by cholesterol. However, aggregated LDL, a major form of LDL in atherosclerotic lesions, can deliver enormous amounts of cholesterol to macrophages and cause foam-cell formation⁶². The likely explanation is that one or more receptors other than LDL receptors are involved and/or that LDL receptor downregulation is not complete in these macrophages. In terms of non-native LDL, a survey of the literature over the past 10–20

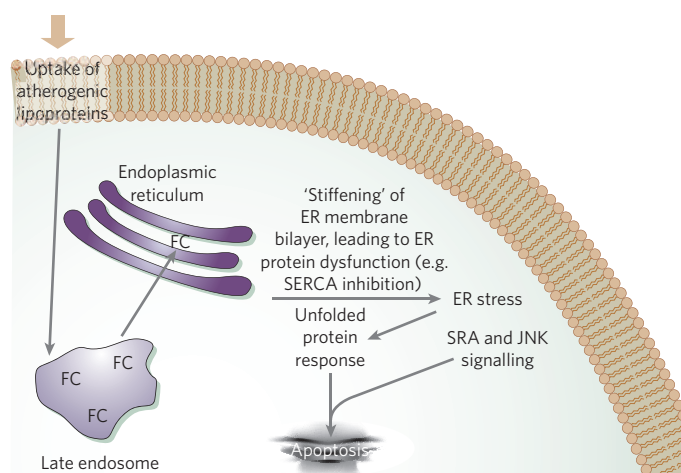


Figure 4 | Free cholesterol-induced apoptosis in macrophages. In macrophages in advanced atherosclerotic lesions, the internalization of atherogenic lipoproteins leads to FC accumulation, perhaps due to defective ACAT and/or overactive neutral CE hydrolase. When the FC:phospholipid ratio in the ER membrane reaches a certain level, integral ER membrane proteins such as sarco(endo)plasmic reticulum ATPase (SERCA) become inactive. Inactivation of SERCA corresponds closely to an increase in order parameter ('stiffening') of the bilayer. Perturbation of ER function by this mechanism leads to activation of the unfolded protein response and other ER stress pathways, which, together with signalling involving the type A scavenger receptor, triggers apoptosis.

years gives the impression that foam cells are formed mostly, if not exclusively, by the uptake of oxidized LDL. However, most forms of oxidized LDL are not particularly potent inducers of foam-cell formation in cultured macrophages⁶². One explanation for this often-neglected finding is that oxidized LDL-derived cholesterol is poorly trafficked from late endosomes to ACAT in the ER⁶². *In vivo*, there are studies showing the importance of two oxidized LDL receptors — the type A scavenger receptor and CD36 (refs 63, 64) — but these findings have been questioned by a recent study in mice⁶⁵. Moreover, the most reliable anti-oxidant trials in humans have not shown a benefit of vitamin E or other anti-oxidants in decreasing the incidence of atherosclerotic heart disease⁶⁶. Finally, a type of atherogenic lipoprotein often neglected in the discussion of foam-cell formation is the remnant lipoprotein class⁶⁷. Remnant lipoproteins result from the partial catabolism and subsequent cholesterol enrichment of triglyceride-rich lipoproteins made by enterocytes and hepatocytes. Remnant lipoproteins are avidly internalized by cultured macrophages and are potent inducers of ACAT activation and foam-cell formation. Moreover, they are abundant in the intima of atherosclerotic lesions, and their levels in plasma are strongly associated with the presence of foam cells and incidence of atherosclerotic vascular disease in animal models and humans⁶⁷. In summary, a number of 'atherogenic' lipoproteins may cause macrophage foam-cell formation during early atherogenesis, and it is almost certain that a combination of these lipoproteins carry out this role *in vivo*.

Regarding the functions of foam cells in atherogenesis, studies with genetically altered mice have uniformly demonstrated the pro-atherogenic role of macrophage foam cells in early lesions⁵⁵. For example, the lesion area is substantially decreased in mice with defective macrophage development resulting from absent M-CSF or in mice with perturbed monocyte chemokines or chemokine receptors⁵⁵. Similar results are found when early lesional macrophages are depleted by enhanced apoptosis⁶⁸. Although the mechanisms of foam-cell-induced atherogenicity are not known, the ability of foam cells to secrete inflammatory cytokines and matrix metalloproteinases are likely to be contributing factors^{55,69}. Macrophage foam cells may also participate in other early atherogenic processes, such as smooth-muscle-cell migration and T-cell-mediated inflammatory and immune responses⁶⁹. A major question is whether these and other roles of foam cells in early atherogenesis are specifically induced by cholesterol loading *per se* or whether cholesterol loading represents a parallel, non-causative event in macrophage-mediated early atherogenesis. Surprisingly, there is a paucity of data addressing this fundamental question.

Late stages of atherosclerosis

Early atherosclerotic lesions are not symptomatic because arterial lumen occlusion is not great enough to compromise blood flow⁷⁰. This lack of occlusion is aided by outward remodelling of the affected region of the arterial wall. After years of gradual lesion development, foam cells, smooth muscle cells, extracellular matrix material and smooth-muscle-cell-derived scar tissue can lead to slowly progressive lumen occlusion, but symptoms are usually absent because organ blood flow is restored by compensatory, hypoxia-driven neovascularization. If this compensatory process becomes compromised, the patient may experience stable, exercise-induced compromise of blood flow (for example, exercise-induced angina) but not acute cardiovascular events⁷⁰. Importantly, the smooth-muscle-cell-derived scar tissue forms a fibrous cap that covers and essentially 'protects' the underlying lesion, and these lesions tend to be relatively stable⁷¹.

A minority of lesions progress to a point in which they precipitate acute vascular events, including sudden death, acute myocardial infarction, unstable angina or ischaemic stroke⁷¹. These events are caused by acute, occlusive luminal thrombosis, which, because of the suddenness of lumen occlusion, leads to organ damage. This process occurs over minutes, so there is not enough time for compensatory responses. Pathological observations of affected arteries in patients

suffering from acute events has led to the plaque-disruption theory of acute atherothrombosis^{71–74}. According to this theory, a minority of plaques become necrotic and highly inflammatory, which eventually leads to breakdown of the protective fibrous cap or to erosion of the endothelial cell layer. These events, in turn, expose the luminal blood to underlying plaque material, which promotes coagulation and thrombosis. Of interest, these rare events do not necessarily occur in the largest plaques, but rather those that have large areas of necrosis.

What promotes plaque disruption? According to one theory, late lesional macrophages secrete matrix metalloproteinases, and these enzymes lead to breakdown of the fibrous cap⁷⁵. *In vitro* studies have supported this idea and have suggested that inflammatory mediators promote macrophages to secrete the proteases. However, definitive *in vivo* data for this idea are lacking. Another theory proposes that death of smooth muscle cells promotes plaque instability because intimal smooth muscle cells synthesize the collagen that makes up the protective cap⁷⁶. A third theory holds that macrophage death is important, because it is this event, in the absence of efficient phagocytic clearance of apoptotic cells, that gives rise to the necrotic core⁷⁷. In this regard, there is evidence for defective phagocytic clearance of apoptotic macrophages in advanced atherosclerotic lesions, which leads to post-apoptotic necrosis of the cells^{78,79}. By contrast, phagocytic clearance of apoptotic macrophages seems to be intact in early lesions⁷⁹. As mentioned above, there are very strong spatial and temporal correlations between necrotic cores and plaque disruption. Although direct causality has not yet been proven *in vivo*, necrotic cores are rich in proteases, inflammatory molecules and pro-coagulation and thrombosis factors^{74,76,79}.

There are a number of theories to explain late lesional macrophage death, including exposure to oxysterols, deprivation of growth factors, interaction with cytotoxic cytokines and intracellular accumulation of excess unesterified or 'free' cholesterol (FC)^{80,81}. Support for the latter mechanism comes from *in vivo* studies showing that late lesional macrophages accumulate large amounts of FC and from *in vitro* studies showing that FC accumulation is a potent inducer of macrophage apoptosis³¹. It is not known why late lesional macrophages accumulate FC, but it is likely to be due to a combination of perturbed cholesterol esterification and diminished cholesterol efflux. Direct proof for dysfunctional ACAT or efflux proteins (such as ABCA1 and ABCG1) in late lesions is lacking. However, *in vivo* observations are strongly consistent with the idea of a dysfunctional ACAT pathway (below).

Mechanistic studies have begun to reveal a fascinating series of signal-transduction pathways that mediate FC-induced macrophage death³¹ (Fig. 4). The key initiating event is trafficking of lipoprotein-derived FC to the ER membrane bilayer, which normally has a low cholesterol:phospholipid ratio and is therefore relatively disordered. Upon enrichment with FC, the order parameter of the ER membrane increases, and this increase is very closely correlated with loss of activity of an integral ER membrane protein, sarco(endo)plasmic reticulum ATPase (SERCA)⁸², a protein related to the sarcoplasmic reticulum Ca²⁺-ATPases found in muscle. The significance of this finding is twofold. First, it probably indicates that other integral membrane proteins in the ER become dysfunctional in FC-loaded macrophages. Second, loss of SERCA function would be expected to result in depleted ER calcium stores. Indeed, careful measurements have shown that ER calcium pools are depleted within about 2 hours of FC loading, and this event may at least trigger subsequent cellular events⁸³. As discussed above, the activity of the SERCA pumps can be affected by changes in bilayer properties, such as thickness, which would be increased upon cholesterol loading.

Within 5 hours of FC loading, upstream and downstream molecules in the ER stress pathway known as the unfolded protein response (UPR) are activated⁸³. As alluded to above, one contributing factor could be depletion of ER calcium stores, which renders chaperones dysfunctional and thereby triggers activation of the UPR. However, the dysfunction of other ER membrane proteins might also contribute to UPR activation. One of several UPR effector proteins is a protein

called CHOP (GADD153), which can, in turn, affect the transcription of a number of genes that participate in the UPR programme⁸⁴. In other systems, CHOP can participate in an apoptosis response, and FC-induced apoptosis is markedly inhibited in Chop^{-/-} macrophages⁸³. *In vivo* evidence for these events includes the demonstration that apoptotic macrophages surrounding necrotic cores in advanced lesions are filled with FC and show evidence of UPR activation^{83,85,86}.

Most importantly, macrophages with a heterozygous mutation in the NPC1 have a defect in trafficking of lipoprotein-cholesterol to the ER, and, as expected, this inhibits FC-induced UPR activation and apoptosis despite increased cellular stores of FC^{83,85}. Analysis of advanced lesions from *Npc1*^{+/-}; *ApoE*^{-/-} mice revealed a marked decrease in late lesional necrosis and macrophage apoptosis compared with similarly sized late lesions of *ApoE*^{-/-} mice⁸⁵. Finally, another consequence of FC-induced ER stress in macrophages is activation of mitogen-activated protein kinase pathways and NFκB⁸⁷. Among the consequences of these events is marked secretion of tumour necrosis factor α and interleukin-6, two inflammatory cytokines that are thought to play important roles in late lesional plaque disruption.

In summary, a series of cell-biological events in a subset of advanced atherosclerotic lesions leads to plaque instability, which, in turn, precipitates acute thrombosis and vascular occlusion. Among these events are secretion of proteases and inflammatory cytokines by macrophages and death of macrophages and smooth muscle cells in the setting of defective phagocytic clearance. Macrophage death seems to be particularly important, because it is this event that gives rise to the destabilizing necrotic core. Increasing evidence suggests that an important cause of late lesional macrophage-mediated inflammation and macrophage death is the accumulation of excess intracellular FC. New studies have revealed a number of signal-transduction pathways, centred on the ER, that account for these cellular effects of FC. Novel therapeutic strategies based on this new insight may provide the means to prevent plaque destabilization and acute atherothrombotic vascular events.

Other diseases

Membrane organization is important for many basic cell functions, and so it would be expected that changes in cholesterol or other aspects of lipid organization have a role in many diseases. There have, for example, been reports for many years that membrane organization and order might be altered in several cancers⁸⁸, but it remains unclear whether such changes play a part in disease progression or are merely byproducts of other metabolic changes. Other reports suggest that statins, which are now among the most widely prescribed drugs, may have uses in cancer chemotherapy⁸⁹, and they may alter endothelial cell function and suppress some inflammatory responses. It is likely that many of these effects are not directly related to their effects on cholesterol but are related to changes in other molecules, such as isoprenoids, that share the same initial biochemical synthetic steps as cholesterol⁹⁰. Isoprenyl groups are important for anchoring several regulatory GTPases such as Ras and Rho in the membrane, and many of these pleiotropic effects of statins may be a consequence of changes in signalling pathways that use these GTPases.

The role of cholesterol and lipids in Alzheimer's disease has been actively studied for over a decade, on the basis of the observation that there is a genetic linkage between age of onset of Alzheimer's disease and the presence of the ε4 allele of apolipoprotein E (ApoE). Polymorphisms in other proteins involved in cholesterol metabolism may also have a genetic linkage with this disease⁷. ApoE is one of the main carriers of cholesterol in the brain, and it seems possible that alterations in cholesterol distribution or levels might have a role in formation of amyloid deposits. The amyloid in Alzheimer's disease is formed by aggregation of a 39–42-residue peptide, the Aβ peptide, which is formed by two proteolytic cleavages of a transmembrane protein, the amyloid precursor protein (APP). These cleavages take place in intracellular organelles. In tissue culture studies, severe lowering of cellular

cholesterol (more than 35% reduction) partly inhibited the formation of the Aβ peptide, but moderate reduction in cellular cholesterol increased the formation of Aβ peptide⁹¹. Furthermore, rodents treated with statins can have increased amyloid production⁹¹, and recent studies indicate that treatment with statins does not reduce the amyloid burden in humans⁷. Nevertheless, statins may be neuroprotective, perhaps because of their pleiotropic effects on endothelial cell function and as suppressors of inflammation^{7,90}. There is still not a good mechanistic explanation for the association of the ε4 allele of ApoE with age of onset of Alzheimer's disease.

Future work

In the past several years there has been increased interest in the role of the lipids in biological membranes. The role of lipids and sterol derivatives as signalling molecules and second messengers is well established, but important new discoveries on the signalling roles of these molecules continue to be made. In this review, we have focused on the more subtle role of lipids and cholesterol in regulating the biophysical properties of membranes and how this affects cell physiology. Recent evidence points to the existence of coexisting microdomains within a single membrane, especially the plasma membrane, even though many important properties of these microdomains remain poorly characterized. These domains are important for regulating some signalling pathways, and we are beginning to understand how this may work in a few cases. Much work needs to be done to better characterize the biophysical properties of biological membranes and the effects that these properties have on membrane proteins.

Cells and organisms have developed extraordinarily sophisticated mechanisms for controlling the lipid composition, and hence the properties, of biological membranes. This control is based on regulating free cholesterol levels and also properties such as the degree of saturation of fatty acids. In atherosclerosis we have one clear example of what goes wrong when these homeostatic regulatory mechanisms are overwhelmed. In both the early and late stages of atherosclerosis there is evidence that changes in membrane bilayer properties influence disease progression. The roles of changes in bilayer properties in other diseases such as Alzheimer's or type II diabetes/metabolic syndrome is less clear, but this may be an area for significant new discoveries of disease mechanisms and treatments. ■

1. Lee, A. G. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta* **1666**, 62–87 (2004).
2. Simons, K. & Vaz, W. L. Model systems, lipid rafts, and cell membranes. *Annu. Rev. Biophys. Biomol. Struct.* **33**, 269–295 (2004).
3. Mukherjee, S. & Maxfield, F. R. Membrane domains. *Annu. Rev. Cell Dev. Biol.* **20**, 839–866 (2004).
4. Holowka, D. *et al.* Lipid segregation and IgE receptor signaling: A decade of progress. *Biochim. Biophys. Acta* doi:10.1016/j.bbamcr.2005.06.007 (2005).
5. Goldstein, J. L. & Brown, M. S. Molecular medicine. The cholesterol quartet. *Science* **292**, 1310–1312 (2001).
6. Anitschkow, N. & Chalotow, S. Über experimentelle cholesterinsteatose und ihre bedeutung für die einiger pathologischer prozesse. *Zentralbl. Allg. Pathol.* **24**, 1–9 (1913).
7. Wolozin, B. Cholesterol, statins and dementia. *Curr. Opin. Lipidol.* **15**, 667–672 (2004).
8. Simons, K. & Ikonen, E. Functional rafts in cell membranes. *Nature* **387**, 569–572 (1997).
9. Munro, S. Lipid rafts: elusive or illusive? *Cell* **115**, 377–388 (2003).
10. Sprong, H., van der Sluijs, P. & van Meer, G. How proteins move lipids and lipids move proteins. *Nature Rev. Mol. Cell Biol.* **2**, 504–513 (2001).
11. Schroeder, R. J., Ahmed, S. N., Zhu, Y., London, E. & Brown, D. A. Cholesterol and sphingolipid enhance the Triton X-100 insolubility of glycosylphosphatidylinositol-anchored proteins by promoting the formation of detergent-insoluble ordered membrane domains. *J. Biol. Chem.* **273**, 1150–1157 (1998).
12. Resh, M. D. Membrane targeting of lipid modified signal transduction proteins. *Subcell. Biochem.* **37**, 217–232 (2004).
13. Wang, T. Y., Leventis, R. & Silvius, J. R. Partitioning of lipidated peptide sequences into liquid-ordered lipid domains in model and biological membranes. *Biochemistry* **40**, 13031–13040 (2001).
14. Melkonian, K. A., Ostermeyer, A. G., Chen, J. Z., Roth, M. G. & Brown, D. A. Role of lipid modifications in targeting proteins to detergent-resistant membrane rafts. Many raft proteins are acylated, while few are prenylated. *J. Biol. Chem.* **274**, 3910–3917 (1999).
15. Bretscher, M. S. & Munro, S. Cholesterol and the Golgi apparatus. *Science* **261**, 1280–1281 (1993).
16. Mitra, K., Ubarretxena-Belandia, I., Taguchi, T., Warren, G. & Engelman, D. M. Modulation of the bilayer thickness of exocytic pathway membranes by membrane proteins rather than cholesterol. *Proc. Natl Acad. Sci. USA* **101**, 4083–4038 (2004).
17. Pierini, L. M. *et al.* Membrane lipid organization is critical for human neutrophil polarization. *J. Biol. Chem.* **278**, 10831–10841 (2003).

18. Feramisco, J. D. *et al.* Intramembrane aspartic acid in SCAP protein governs cholesterol-induced conformational change. *Proc. Natl Acad. Sci. USA* **102**, 3242–3247 (2005).
19. Hao, M. *et al.* Vesicular and non-vesicular sterol transport in living cells. The endocytic recycling compartment is a major sterol storage organelle. *J. Biol. Chem.* **277**, 609–617 (2002).
20. Maxfield, F. R. & Wustner, D. Intracellular cholesterol transport. *J. Clin. Invest.* **110**, 891–898 (2002).
21. Vainio, S. & Ikonen, E. Macrophage cholesterol transport: a critical player in foam cell formation. *Ann. Med.* **35**, 146–155 (2003).
22. Soccio, R. E. & Breslow, J. L. Intracellular cholesterol transport. *Arterioscler. Thromb. Vasc. Biol.* **24**, 1150–1160 (2004).
23. Strauss, J. F., Kishida, T., Christenson, L. K., Fujimoto, T. & Hiroi, H. START domain proteins and the intracellular trafficking of cholesterol in steroidogenic cells. *Mol. Cell Endocrinol.* **202**, 59–65 (2003).
24. Wustner, D., Mondal, M., Tabas, I. & Maxfield, F. R. Direct observation of rapid internalization and intracellular transport of sterol by macrophage foam cells. *Traffic* **6**, 396–412 (2005).
25. Sleat, D. E. *et al.* Genetic evidence for nonredundant functional cooperativity between NPC1 and NPC2 in lipid transport. *Proc. Natl Acad. Sci. USA* **101**, 5886–5891 (2004).
26. Carstea, E. *et al.* Niemann–Pick C1 disease gene: homology to mediators of cholesterol homeostasis. *Science* **277**, 228–231 (1997).
27. Blanchette-Mackie, E. Intracellular cholesterol trafficking: role of the NPC1 protein. *Biochim. Biophys. Acta* **1486**, 171–183 (2000).
28. Mukherjee, S. & Maxfield, F. R. Lipid and cholesterol trafficking in NPC. *Biochim. Biophys. Acta* **1685**, 28–37 (2004).
29. Chang, T. Y., Chang, C. C. & Cheng, D. Acyl-coenzyme A:cholesterol acyltransferase. *Annu. Rev. Biochem.* **66**, 613–638 (1997).
30. Yeaman, S. J. Hormone-sensitive lipase—new roles for an old enzyme. *Biochem. J.* **379**, 11–22 (2004).
31. Tabas, I. Consequences of cellular cholesterol accumulation: basic concepts and physiological implications. *J. Clin. Invest.* **110**, 905–911 (2002).
32. Zhang, Y. *et al.* Cholesterol is superior to 7-ketocholesterol or 7 α -hydroxycholesterol as an allosteric activator for acyl-coenzyme A:cholesterol acyltransferase 1. *J. Biol. Chem.* **278**, 11642–11647 (2003).
33. Lange, Y., Ye, J. & Steck, T. L. How cholesterol homeostasis is regulated by plasma membrane cholesterol in excess of phospholipids. *Proc. Natl Acad. Sci. USA* **101**, 11664–11667 (2004).
34. Brown, M. S. & Goldstein, J. L. A proteolytic pathway that controls the cholesterol content of membranes, cells, and blood. *Proc. Natl Acad. Sci. USA* **96**, 11041–11048 (1999).
35. Tall, A. R., Costet, P. & Wang, N. Regulation and mechanisms of macrophage cholesterol efflux. *J. Clin. Invest.* **110**, 899–904 (2002).
36. Pohl, A., Devaux, P. F. & Herrmann, A. Function of prokaryotic and eukaryotic ABC proteins in lipid transport. *Biochim. Biophys. Acta* **1733**, 29–52 (2005).
37. Fielding, C. J. & Fielding, P. E. Cellular cholesterol efflux. *Biochim. Biophys. Acta* **1533**, 175–189 (2001).
38. Sun, Y. *et al.* Stearoyl-CoA desaturase inhibits ATP-binding cassette transporter A1-mediated cholesterol efflux and modulates membrane domain structure. *J. Biol. Chem.* **278**, 5813–5820 (2003).
39. Wang, N. & Tall, A. R. Regulation and mechanisms of ATP-binding cassette transporter A1-mediated cellular cholesterol efflux. *Arterioscler. Thromb. Vasc. Biol.* **23**, 1178–1184 (2003).
40. Schroefer, G. Oxysterols: modulators of cholesterol metabolism and other processes. *Physiol. Rev.* **80**, 361–554 (2000).
41. Costet, P. *et al.* Retinoic acid receptor-mediated induction of ABCA1 in macrophages. *Mol. Cell Biol.* **23**, 7756–7766 (2003).
42. Chawla, A., Repa, J. J., Evans, R. M. & Mangelsdorf, D. J. Nuclear receptors and lipid physiology: opening the X-files. *Science* **294**, 1866–1870 (2001).
43. Oram, J. F. ATP-binding cassette transporter A1 and cholesterol trafficking. *Curr. Opin. Lipidol.* **13**, 373–381 (2002).
44. Horton, J. D., Goldstein, J. L. & Brown, M. S. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin. Invest.* **109**, 1125–1131 (2002).
45. Sampath, H. & Ntambi, J. M. Polyunsaturated fatty acid regulation of genes of lipid metabolism. *Annu. Rev. Nutr.* **25**, 317–340 (2005).
46. Kolter, T. & Sandhoff, K. Principles of lysosomal membrane digestion—stimulation of sphingolipid degradation by sphingolipid activator proteins and anionic lysosomal lipids. *Annu. Rev. Cell Dev. Biol.* **21**, 81–103 (2004).
47. Kobayashi, T. *et al.* Late endosomal membranes rich in lysobisphosphatidic acid regulate cholesterol transport. *Nature Cell Biol.* **1**, 113–118 (1999).
48. Sturley, S. L., Patterson, M. C., Balch, W. & Liscum, L. The pathophysiology and mechanisms of NP-C disease. *Biochim. Biophys. Acta* **1685**, 83–87 (2004).
49. Friedland, N., Liou, H. L., Lobel, P. & Stock, A. M. Structure of a cholesterol-binding protein deficient in Niemann–Pick type C2 disease. *Proc. Natl Acad. Sci. USA* **100**, 2512–2517 (2003).
50. Pagano, R. E. Endocytic trafficking of glycosphingolipids in sphingolipid storage diseases. *Phil. Trans. R. Soc. Lond. B* **358**, 885–891 (2003).
51. Stefkova, J., Poledne, R. & Hubacek, J. A. ATP-binding cassette (ABC) transporters in human metabolism and diseases. *Physiol. Res.* **53**, 235–243 (2004).
52. Williams, K. J. & Tabas, I. The response-to-retention hypothesis of early atherogenesis. *Arterioscler. Thromb. Vasc. Biol.* **15**, 551–561 (1995).
53. Hansson, G. K. Inflammation, atherosclerosis, and coronary artery disease. *N. Engl. J. Med.* **352**, 1685–1695 (2005).
54. Berliner, J. A., Subbanagounder, G., Leitinger, N., Watson, A. D. & Vora, D. Evidence for a role of phospholipid oxidation products in atherogenesis. *Trends Cardiovasc. Med.* **11**, 142–147 (2001).
55. Glass, C. K. & Witztum, J. L. Atherosclerosis. The road ahead. *Cell* **104**, 503–516 (2001).
56. Brown, M. S. & Goldstein, J. L. Lipoprotein metabolism in the macrophage: implications for cholesterol deposition in atherosclerosis. *Annu. Rev. Biochem.* **52**, 223–261 (1983).
57. Kruth, H. S. *et al.* Macropinocytosis is the endocytic pathway that mediates macrophage foam cell formation with native low density lipoprotein. *J. Biol. Chem.* **280**, 2352–2360 (2005).
58. Buton, X. *et al.* Unique cellular events occurring during the initial interaction of macrophages with matrix-retained or methylated aggregated low density lipoprotein (LDL). Prolonged cell-surface contact during which LDL-cholesterol ester hydrolysis exceeds LDL protein degradation. *J. Biol. Chem.* **274**, 32112–32121 (1999).
59. Sakr, S. W. *et al.* The uptake and degradation of matrix-bound lipoproteins by macrophages require an intact actin cytoskeleton, Rho family GTPases, and myosin ATPase activity. *J. Biol. Chem.* **276**, 37649–37658 (2001).
60. Qin, C., Nagao, T., Grosheva, I., Maxfield, F. R. & Pierini, L. M. Elevated plasma membrane cholesterol content alters macrophage signaling and function. *Arterioscler. Thromb. Vasc. Biol.* (in press).
61. Gerrity, R. G. The role of the monocyte in atherogenesis: I. Transition of blood-borne monocytes into foam cells in fatty lesions. *Am. J. Pathol.* **103**, 181–190 (1981).
62. Tabas, I. Nonoxidative modifications of lipoproteins in atherogenesis. *Annu. Rev. Nutr.* **19**, 123–139 (1999).
63. Suzuki, H. *et al.* A role for macrophage scavenger receptors in atherosclerosis and susceptibility to infection. *Nature* **386**, 292–296 (1997).
64. Febbraio, M., Hajjar, D. P. & Silverstein, R. L. CD36: a class B scavenger receptor involved in angiogenesis, atherosclerosis, inflammation, and lipid metabolism. *J. Clin. Invest.* **108**, 785–791 (2001).
65. Moore, K. J. *et al.* Loss of receptor-mediated lipid uptake via scavenger receptor A or CD36 pathways does not ameliorate atherosclerosis in hyperlipidemic mice. *J. Clin. Invest.* **115**, 2192–2201 (2005).
66. Williams, K. J. & Fisher, E. A. Oxidation, lipoproteins, and atherosclerosis: which is wrong, the antioxidants or the theory? *Curr. Opin. Clin. Nutr. Metab. Care* **8**, 139–146 (2005).
67. Twickler, T., Dallinga-Thie, G. M., Chapman, M. J. & Cohn, J. S. Remnant lipoproteins and atherosclerosis. *Curr. Atheroscler. Rep.* **7**, 140–147 (2005).
68. Arai, S. *et al.* A role for the apoptosis inhibitory factor AIM/Spa/Api6 in atherosclerosis development. *Cell Metabolism* **1**, 201–213 (2005).
69. Libby, P. & Clinton, S. K. The role of macrophages in atherosclerosis. *Curr. Opin. Lipidol.* **4**, 355–363 (1993).
70. Abrams, J. Clinical practice. Chronic stable angina. *N. Engl. J. Med.* **352**, 2524–2533 (2005).
71. Aikawa, M. & Libby, P. The vulnerable atherosclerotic plaque: pathogenesis and therapeutic approach. *Cardiovasc. Pathol.* **13**, 125–138 (2004).
72. Shah, P. K. Insights into the molecular mechanisms of plaque rupture and thrombosis. *Indian Heart J.* **57**, 21–30 (2005).
73. Corti, R., Hutter, R., Badimon, J. J. & Fuster, V. Evolving concepts in the triad of atherosclerosis, inflammation and thrombosis. *J. Thromb. Thrombolysis* **17**, 35–44 (2004).
74. Virmani, R., Burke, A. P., Kolodgie, F. D. & Farb, A. Vulnerable plaque: the pathology of unstable coronary lesions. *J. Interv. Cardiol.* **15**, 439–446 (2002).
75. Libby, P. *et al.* Macrophages and atherosclerotic plaque stability. *Curr. Opin. Lipidol.* **7**, 330–335 (1996).
76. Geng, Y. J. & Libby, P. Progression of atheroma: a struggle between death and procreation. *Arterioscler. Thromb. Vasc. Biol.* **22**, 1370–1380 (2002).
77. Tabas, I. Apoptosis and plaque destabilization in atherosclerosis: the role of macrophage apoptosis induced by cholesterol. *Cell Death Differ.* **11** (Suppl. 1), S12–S16 (2004).
78. Schrijvers, D. M., De Meyer, G. R., Kockx, M. M., Herman, A. G. & Martinet, W. Phagocytosis of apoptotic cells by macrophages is impaired in atherosclerosis. *Arterioscler. Thromb. Vasc. Biol.* **25**, 1256–1261 (2005).
79. Tabas, I. Consequences and therapeutic implications of macrophage apoptosis in atherosclerosis. The importance of lesion stage and phagocytic efficiency. *Arterioscler. Thromb. Vasc. Biol.* **25**, 2255–2264 (2005).
80. Kockx, M. M. Apoptosis in the atherosclerotic plaque: quantitative and qualitative aspects. *Arterioscler. Thromb. Vasc. Biol.* **18**, 1519–1522 (1998).
81. Nhan, T. Q., Liles, W. C. & Schwartz, S. M. Role of caspases in death and survival of the plaque macrophage. *Arterioscler. Thromb. Vasc. Biol.* **25**, 895–903 (2005).
82. Li, Y. *et al.* Enrichment of endoplasmic reticulum with cholesterol inhibits sarcoplasmic-endoplasmic reticulum calcium ATPase-2b activity in parallel with increased order of membrane lipids: implications for depletion of endoplasmic reticulum calcium stores and apoptosis in cholesterol-loaded macrophages. *J. Biol. Chem.* **279**, 37030–37039 (2004).
83. Feng, B. *et al.* The endoplasmic reticulum is the site of cholesterol-induced cytotoxicity in macrophages. *Nature Cell Biol.* **5**, 781–792 (2003).
84. Kaufman, R. J. Orchestrating the unfolded protein response in health and disease. *J. Clin. Invest.* **110**, 1389–1398 (2002).
85. Feng, B. *et al.* Niemann–Pick C heterozygosity confers resistance to lesional necrosis and macrophage apoptosis in murine atherosclerosis. *Proc. Natl Acad. Sci. USA* **100**, 10423–10428 (2003).
86. Zhou, J., Lhotak, S., Hilditch, B. A. & Austin, R. C. Activation of the unfolded protein response occurs at all stages of atherosclerotic lesion development in apolipoprotein E-deficient mice. *Circulation* **111**, 1814–1821 (2005).
87. Li, Y. *et al.* Free cholesterol-loaded macrophages are an abundant source of tumor necrosis factor- α and interleukin-6: Model of NK- κ B- and Map kinase-dependent inflammation in advanced atherosclerosis. *J. Biol. Chem.* **280**, 21763–21772 (2005).
88. Schroeder, F. Fluorescence probes in metastatic B16 melanoma membranes. *Biochim. Biophys. Acta* **776**, 299–312 (1984).
89. Mo, H. & Elson, C. E. Studies of the isoprenoid-mediated inhibition of mevalonate synthesis applied to cancer chemotherapy and chemoprevention. *Exp. Biol. Med.* **229**, 567–585 (2004).
90. Vaughan, C. J. & Gotto, A. M. Jr. Update on statins: 2003. *Circulation* **110**, 886–892 (2004).
91. Abad-Rodriguez, J. *et al.* Neuronal membrane cholesterol loss enhances amyloid peptide generation. *J. Cell Biol.* **167**, 953–960 (2004).
92. London, E. & Brown, D. A. Insolubility of lipids in Triton X-100: physical origin and relationship to sphingolipid/cholesterol membrane domains (rafts). *Biochim. Biophys. Acta* **1508**, 182–195 (2000).
93. Huang, J. & Feigenson, G. W. A microscopic interaction model of maximum solubility of cholesterol in lipid bilayers. *Biophys. J.* **76**, 2142–2157 (1999).
94. McConnell, H. M. & Radhakrishnan, A. Condensed complexes of cholesterol and phospholipids. *Biochim. Biophys. Acta* **1610**, 159–173 (2003).

95. Brown, D. A. & London, E. Structure and origin of ordered lipid domains in biological membranes. *J. Membr. Biol.* **164**, 103–114 (1998).
96. Feigenson, G. W. & Buboltz, J. T. Ternary phase diagram of dipalmitoyl-PC/dilauroyl-PC/cholesterol: nanoscopic domain formation driven by cholesterol. *Biophys. J.* **80**, 2775–2788 (2001).
97. Megha & London, E. Ceramide selectively displaces cholesterol from ordered lipid domains (rafts): implications for lipid raft structure and function. *J. Biol. Chem.* **279**, 9997–10004 (2004).
98. Wang, T. Y. & Silvius, J. R. Sphingolipid partitioning into ordered domains in cholesterol-free and cholesterol-containing lipid bilayers. *Biophys. J.* **84**, 367–378 (2003).
99. Nakamura, K. *et al.* Expression and regulation of multiple murine ATP-binding cassette transporter G1 mRNAs/isoforms that stimulate cellular cholesterol efflux to high density lipoprotein. *J. Biol. Chem.* **279**, 45980–45989 (2004).
100. Wang, N., Lan, D., Chen, W., Matsuura, F. & Tall, A. R. ATP-binding cassette transporters G1 and G4 mediate cellular cholesterol efflux to high-density lipoproteins. *Proc. Natl Acad. Sci. USA* **101**, 9774–9779 (2004).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence should be addressed to F.R.M. (frmaxfie@med.cornell.edu).

Phyllosilicates on Mars and implications for early martian climate

F. Poulet¹, J.-P. Bibring¹, J. F. Mustard², A. Gendrin², N. Mangold³, Y. Langevin¹, R. E. Arvidson⁴, B. Gondet¹, C. Gomez¹ & the Omega Team*

The recent identification of large deposits of sulphates by remote sensing and *in situ* observations has been considered evidence of the past presence of liquid water on Mars. Here we report the unambiguous detection of diverse phyllosilicates, a family of aqueous alteration products, on the basis of observations by the OMEGA imaging spectrometer on board the Mars Express spacecraft. These minerals are mainly associated with Noachian outcrops, which is consistent with an early active hydrological system, sustaining the long-term contact of igneous minerals with liquid water. We infer that the two main families of hydrated alteration products detected—phyllosilicates and sulphates—result from different formation processes. These occurred during two distinct climatic episodes: an early Noachian Mars, resulting in the formation of hydrated silicates, followed by a more acidic environment, in which sulphates formed.

The presence of hydrated minerals on Mars provides a record of water-related processes. Hydrated sulphates have been observed with the OMEGA (Observatoire pour la Mineralogie, l'Eau, les Glaces et l'Activité) instrument on board the European Space Agency (ESA) Mars Express mission, in numerous light-toned layered deposits in Valles Marineris, Aram Chaos, and Terra Meridiani¹ and in deposits adjacent to the north polar cap². Observations in Terra Meridiani by the Opportunity rover of a variety of sulphates in layered rocks also require an active hydrologic system to account for these deposits³.

The presence of phyllosilicates on Mars has been previously suggested on the basis of *in situ* elemental analyses by the Viking Landers⁴, the identification of smectites in some SNC (Shergottite–Nakhilite–Chassigny) meteorites⁵, and remote sensing infrared observations^{6–9}. An unambiguous detection of water-bearing phyllosilicates has been reported over large areas¹⁰. Here we present an overview of the detection of phyllosilicates made by OMEGA, and we discuss their geological context inferred from analyses of imaging data. Phyllosilicates represent a very specific family of highly altered

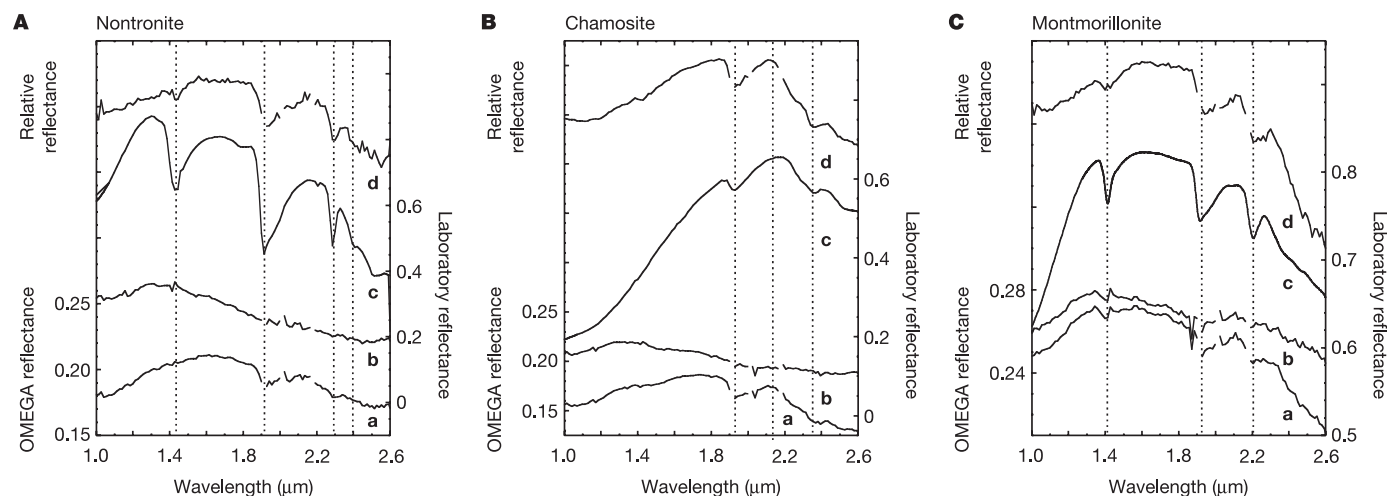


Figure 1 | Phyllosilicate spectra as identified in the OMEGA observations. **A**, Nontronite from the Nili Syrtis Mensae (73.32° E, 29.30° N); **B**, Chamosite from a crater floor in northern Syrtis Major (71.73° E, 17.09° N); **C**, Montmorillonite from light-toned deposits in the vicinity of Mawrth Vallis (20.60° W, 25.53° E). Within each panel, the spectra are labelled as follows: a, I/F atmospherically corrected phyllosilicate spectrum; b, I/F atmospherically corrected reference spectrum; c, laboratory spectra; d, spectral ratio (a divided by b). The OMEGA reflectance scale corresponds to the two lower OMEGA spectra (a and b), the laboratory reflectance to spectrum c, and the relative reflectance to the ratio of OMEGA phyllosilicate/reference spectra (d). The spectra are averages of 5–9 pixels. The reference spectra are selected from regions near to the phyllosilicate spectra, which were acquired during the same OMEGA observation to best-match atmospheric conditions at the time of measurement.

¹Institut d'Astrophysique Spatiale, Université Paris-Sud and CNRS (UMR 8617) F-91405 Orsay, France. ²Geological Sciences, Brown University, Providence, Rhode Island 02912, USA. ³Interactions et Dynamique des Environnements de Surface, F-91405 Orsay, France. ⁴Earth and Planetary Sciences, Washington University, St Louis, Missouri 63130, USA.

*Lists of participants and affiliations appear at the end of the paper

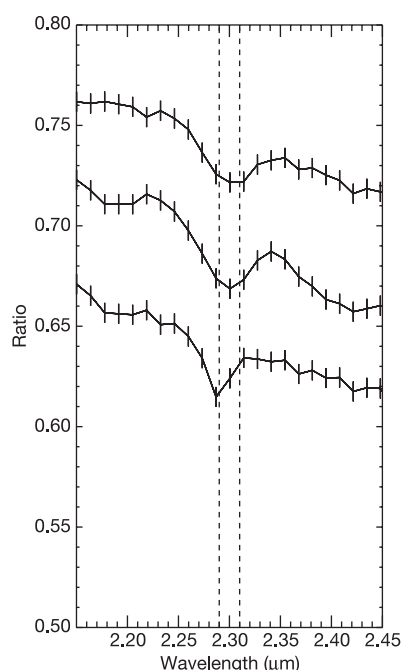


Figure 2 | Variation of the position and shape of the absorption feature in the 2.30- μm region, attributed to varying Fe/Mg abundance. The spectra have been extracted from a terrain in northern Syrtis Major. Three ratioed OMEGA spectra are shown; each spectrum has been divided by a reference spectrum. The ratio process removed instrumental and atmospheric effects. The error bars represent one standard error.

products involving water, so their identification puts constraints on the evolution of Mars. We discuss the implications of our detections to the understanding of the early history of Mars.

Data reduction and mineral identification

The OMEGA instrument acquires three-dimensional (x, y, λ) image cubes, in a spectral domain dominated by solar reflected light (0.3–3.0 μm , with a spectral sampling of ≤ 14 nm), where most minerals exhibit diagnostic absorption bands. The presence of water induces specific vibrational absorptions for the different classes and

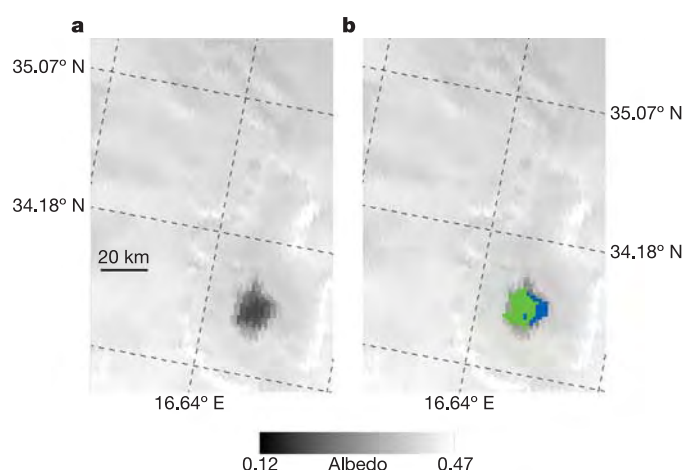


Figure 3 | Dark deposit in a depression in Ismenius Lacus. **a**, Lambertian albedo map at 1.085 μm ; **b**, spatial distributions of pyroxene in green and clays in blue are superimposed for areas exhibiting an absorption feature at 2.15 and 1.93 μm respectively. Fe/Mg clays are present in this dark terrain.

subclasses of hydrated minerals; it is therefore possible to discriminate between major minerals resulting from alteration and aqueous processes, such as carbonates, sulphates, phyllosilicates and zeolites¹.

We applied standard processing and reduction procedures to the OMEGA data in the 1.0–2.6 m wavelength range^{10,11} to identify absorption features due to water of hydration near ~ 1.9 μm and metal–OH vibrations in the 2.2–2.4- μm range. The identification of hydrated silicate is first based on the detection of the 1.9- μm absorption band, calculated using spectral channels at 1.93 μm for the band centre and at 1.86 and 2.14 μm for the continuum. These two wavelengths are selected to lie outside the main atmospheric bands. The 1.9- μm band is due to water molecules, either physically or chemically adsorbed. The latter occurs in phyllosilicates, in which water is bound to interlayer cations¹² or cations in tetrahedral sites¹³. Once hydrated regions are identified by this 1.9- μm band, we focus our analysis on the 2.2–2.4- μm interval. Mineralogical assignment is then based on comparison with laboratory spectra. To enhance the spectral signatures of the materials, we perform spectral ratios. Spectra

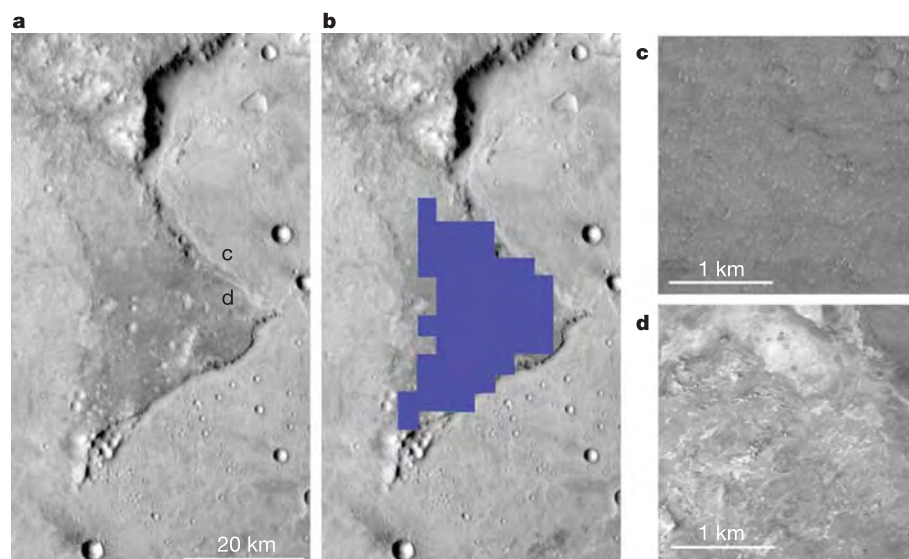


Figure 4 | Detection of Fe-rich clays over Noachian outcrop in Syrtis Major. **a**, THEMIS daytime infrared I02469002 image. The lavas appear brighter than the peninsula outcrop because of the higher temperature resulting from the dark albedo. **b**, Superimposed on the black and white

THEMIS image is the mapping in blue of the Fe-rich clays. The difference in morphology between the early Hesperian pyroxene-rich lava and the unburied Noachian hydrated outcrop can be seen on small portions **c** (dark lava) and **d** (lighter outcrop) of the MOC M0701150 image.

with a 1.9- μm band, exhibiting no features in the 2.2–2.4- μm interval, have been identified, but they are not considered in the present study, because many hydrated minerals exhibit a 1.9- μm band. It is generally considered that absorptions in the 2.2–2.4- μm wavelength region correspond to metal–OH vibrations, where the precise position of this band is a function of the cation species: Al–OH, Fe–OH and Mg–OH vibrational features are centred at 2.2, 2.29 and >2.30 μm respectively^{14–16}. Features in the 2.3–2.35- μm region can also be found in laboratory spectra due to an Al–OH vibration, but they are always associated with a 2.2- μm feature and their strength is much weaker than the 2.2- μm absorption¹⁵.

Figure 1 illustrates the diversity of 1–2.6- μm OMEGA spectra, attributed to phyllosilicates on the basis of their narrow absorption features at ~ 1.9 μm and in the range 2.2–2.4 μm . In Fig. 1A, the spectral features at 1.41, 2.29 and 2.40 μm and the global shape are typical of Fe-rich smectites such as nontronite¹⁷. In Fig. 1B, the spectral features at 1.41 and 2.35 μm and the global shape are those of chamosite, a (Fe/Mg)-phyllosilicate. In Fig. 1C, a third spectral type is identified, with band centres at 1.41, 2.21 and 2.35 μm . Al-rich phyllosilicates such as montmorillonite provide a very good match. Figure 2 shows expanded OMEGA spectra in the 2.3- μm region, to illustrate the degree of sensitivity observed for the position and shape of this feature. Laboratory spectra of smectites have shown a shift from Fe–OH near 2.29 μm to Mg–OH near 2.31 μm (refs 14, 16). The variety of spectra observed in OMEGA data indicates a range of clay composition, from Fe-rich to Mg-rich smectites. We note that no serpentine clay, usually characterized by a strong 2.33–2.34 μm feature, has been detected so far.

Mapping

Maps of phyllosilicates have been built using spectra in which the 1.9- μm band depth exceeds 2%. The variation of the band depth depends on the degree of hydration, the grain size and the relative abundance of phyllosilicates. We do not discuss the quantitative evaluation of the abundances here, because it requires a complex and nonlinear spectral deconvolution taking into account the potential presence of spectrally neutral species.

After 18 months of operation, OMEGA has covered over 75% of the surface of Mars at a 1.5–4.8 km pixel^{−1} sampling. A major outcome of the present work is that phyllosilicates are detected in only a very restricted number of areas, commonly in association with two types of terrains: dark deposits and eroded outcrops. The key regions of each class are discussed separately below.

The dark deposits are mainly located in and around Arabia Terra, northern Syrtis Major, northern Terra Meridiani, and a few small spots are also found in the Xanthe Terra and Lunae Planum regions. The absorption features at 1.9 μm and 2.30 ± 0.01 μm indicate the presence in these deposits of Fe/Mg-smectites. No montmorillonite-like phyllosilicates have been detected in these dark terrains. A typical example of clay-rich dark soil is shown in Fig. 3. The clays are inside a depression but constrained to a part of the dark deposit only; spectral signatures of pyroxenes dominate the rest of the dark surface. The Mars Orbiter Camera (MOC) narrow-angle images show that the hydrated dark material is probably constituted by a thin surface layer of dark material. This region is representative of most of the observed dark clay-rich occurrences. The low albedo seems to indicate that clays, usually bright in terrestrial analogues, are mixed with an opaque material, which does not present diagnostic spectral features in this wavelength range (a ‘spectrally neutral’ component). Two scenarios could be proposed to account for the existence of these clay-rich deposits. The first one is the recent surface alteration of mafic material. However, this should have led to a planet-wide distribution of altered surface material, which is not the case. In the other scenario, the alteration could have taken place much earlier. The altered material would then be buried, and eventually exposed by erosion in very specific locations. The dark deposits would then originate from the local erosion of ancient clay-rich sub-surface

terrains. In some cases, MOC narrow-angle images show the presence of some layered terrains underlying the dark dust, which would favour this second scenario.

The second major type of clay-rich terrains consists of outcrops. Such terrains were first detected in the Syrtis Major region¹⁰. More recently, large occurrences were observed in Nili Fossae and Mawrth Vallis. A few spots have been also identified south of the Isidis basin, northern Hellas, and around Terra Meridiani. The composition of these areas is more diverse than that of the dark deposits. An example of clays in Syrtis Major is given on Fig. 4. The presence of Fe-rich clays perfectly matches the contours of a peninsula of ancient basement; the surrounding younger Syrtis pyroxene-bearing lava flows do not contain hydrated silicates. This confirms our initial interpretation¹⁰

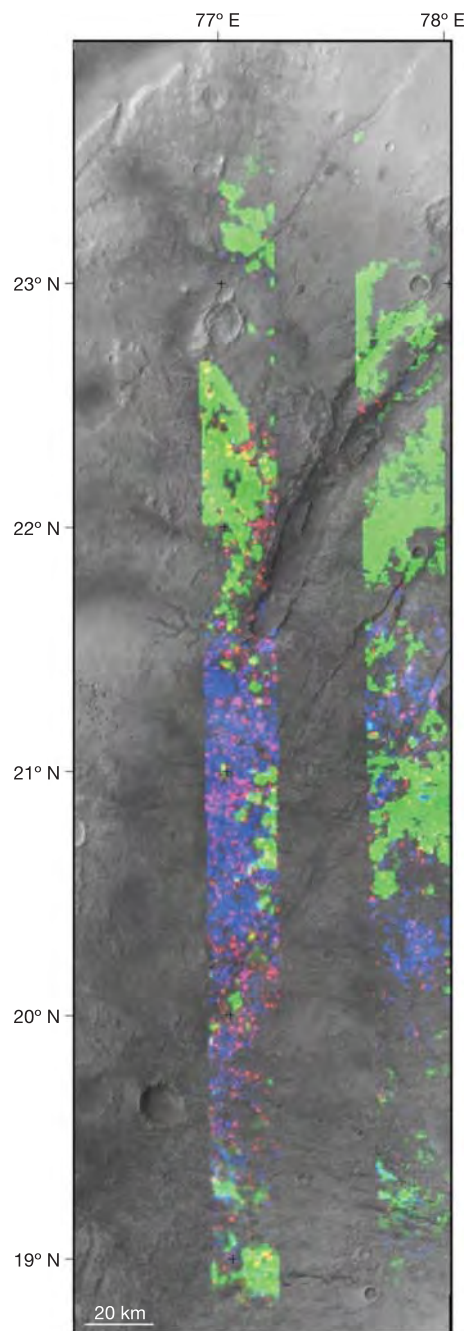


Figure 5 | Spatial distributions of minerals in the Nili Fossae region. Superimposed on a HRSC/Mars Express image, olivine-rich mineral is mapped in green, hydrated minerals are identified by the 1.93- μm band (see text for the definition of the spectral parameter) in purple, and (Fe/Mg)-rich clays are identified by a 2.30- μm feature in red.

that the alteration that produced these clays occurred before the lava outflows, which are dated to the Early Hesperian period¹⁸. Additional evidence for the presence of clays in the crust of this region comes from the detection of clays in the ejecta of several craters. Although some clays could be formed by the impact, provided water was abundant in the impacted material, it is more likely that most, if not all, of the clays have been excavated from altered crust, as shown by clay-rich outcrops outside the ejecta.

Nili Fossae is an interesting area of Mars in which large abundances of olivine have been detected by the Thermal Emission Spectrometer (TES)¹⁹, the Thermal Emission Imaging System (THEMIS)^{20,21} and OMEGA¹¹. The identification of olivine in this region as well as in several areas of Mars has strengthened the cold and dry Mars scenario^{19,20}. In Nili Fossae, OMEGA has also mapped phyllosilicate minerals (Fig. 5). The areas of high olivine abundance are spatially distinct from those containing high concentrations of hydrated minerals. Nili Fossae cratered terrains are Noachian; they experienced strong erosional processes indicated by dissected terrains with rough chaotic texture, isolated mesas and partially eroded craters. Close examination of the high-resolution High Resolution Stereo Camera (HRSC), MOC and THEMIS imaging data shows that the olivine outcrops rest above the clay-bearing substrate. Thus, the olivine-rich rocks appear to have deposited on top of an older, aqueously altered crust. In addition to the various hypotheses accounting for the origin of olivine-rich terrains²¹, they could have originated from the impact that formed the Isidis basin, placing the aqueous-clay-formation episode in the very early Noachian epoch. Alternatively, one may consider that the initial olivine bedrocks were altered to clays, and brought to the surface in a few patches by erosional processes.

OMEGA has discovered a sizeable accumulation of clays in the Mawrth Vallis region between 20° and 28°N and 17° and 22°W (Fig. 6). Phyllosilicates are identified in several light-toned outcrop terrains in the flanks of Mawrth Vallis between -3,200 and -2,700 m as well as on the plateau at an altitude of about -2,300 m. No hydrated mineral has been detected in the valley itself, except for a deposit in a small eroded basin. A remarkable montmorillonite-rich deposit is detected in the more western part (Fig. 6b). THEMIS night-time observations of these terrains are very homogeneous and among the warmest in the region; MOC images indicate geomorphic features typical of eroded terrains. A series of layered deposits occur in these clay-rich areas giving the attributes of material emplaced as sediment²². On the west part of the valley (18°W, 23°N), clays are

identified in the ejecta of a crater, which reinforces the fact that the clays are in the bulk component of the outcrops and were formed before the occurrence of the water discharge that sculpted Mawrth Vallis. The clay-rich areas are surrounded by low-albedo cratered terrains extending to the south towards Terra Meridiani and showing strong signatures of pyroxene. This region is considered to be ancient, with a late Noachian age^{23,24}. Therefore, the presence of clays in eroded outcrops distributed vertically over 500 m provides additional strong evidence for major alteration processes involving liquid water occurring during the early history of Mars. Furthermore, the detection of phyllosilicates in small areas of Arabia Terra and northern Terra Meridiani suggests that the alteration processes could have been intense over this entire region.

Implications for early Mars

From these observations, the presence of clays in outcrops and ejecta strongly supports the following conclusions: (1) the deposits in the crust (Syrtis Major, Nili Fossae) predate the volcanism of Syrtis Major, and possibly the formation of Isidis basin itself; in Mawrth Vallis, clay deposits predate the Noachian/early Hesperian cratering; (2) the clays are a bulk component of the deposits rather than a surface coating or dust layer; (3) the diversity of the composition indicates that the alteration processes affected the variety of igneous rocks (mafic and Al-rich silicates) constituting the martian crust. Note that these conclusions on clay-rich outcrops are consistent with the idea that the dark clay-rich deposits (discussed above) are formed through the erosion of ancient subsurface clay-rich layers.

The formation of clays is controlled by bedrock composition and topography, climate-derived parameters (temperature and long-term availability of liquid water, at or below the surface), availability of water, time and kinetics of mineral reactions²⁵. In terms of bedrock origin, Fe-rich smectites such as nontronites are typical of the alteration of mafic material such as gabbros or basalts²⁶, which are common on Mars. The Al-rich phyllosilicate could either indicate a higher alteration²⁷ or originate from the alteration of more acidic crustal rocks containing Al-rich minerals such as orthoclase²⁶.

In terms of surface temperature and long-term availability of water, the formation of clays, and specifically of smectites, requires conditions very different from those currently observed for Mars²⁸. For clay formation on the Earth, smectites dominate in the areas of moderate alteration of the temperate regions²⁹, rocks of tropical zones are mainly weathered to kaolinite and hydroxides, and rocks of arid regions (<200 mm yr⁻¹ precipitation) or polar regions alter to poorly

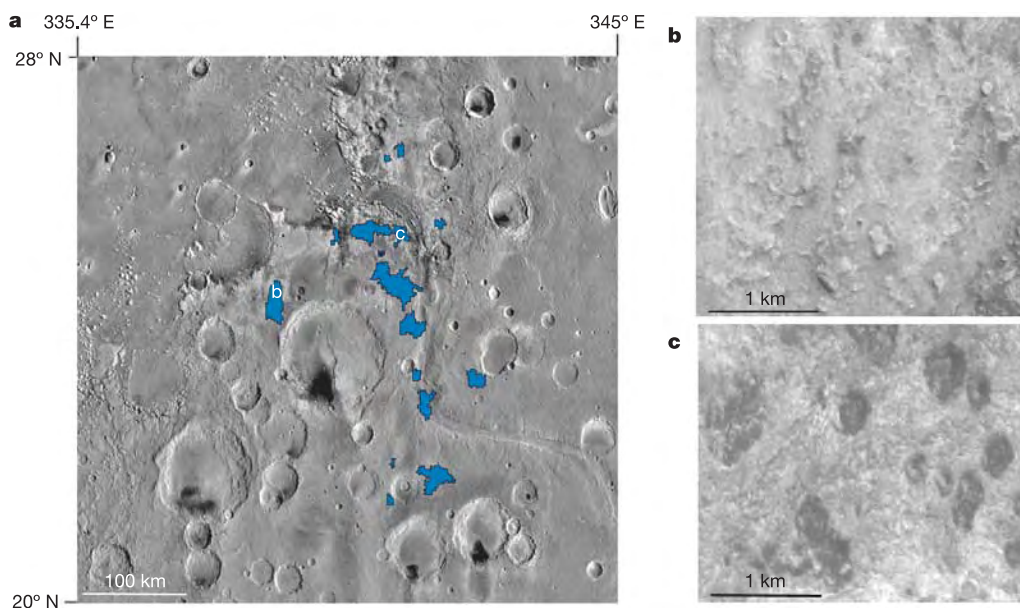


Figure 6 | Identification of clays in Mawrth Vallis. **a**, Map of hydrated minerals in blue over Viking image. Al/Fe/Mg-clays are present. **b**, MOC illustration (E1101550) of the morphology of a montmorillonite-rich outcrop. **c**, MOC illustration (MOC R0801755) of the morphology of a Fe-smectite-rich outcrop. The presence of small mesas indicates strong erosion.

crystalline clays or immature weathering products. In the Dry Valleys of Antarctica, clays are indeed found, as a result of seasonal processes over million of years of activity³⁰. However, a similar mechanism on Mars would result in small amounts of clays as thin coatings and does not account for the clay-rich units detected by OMEGA.

The presence of widespread smectites in the martian Noachian rocks suggests an early active hydrologic system that could have sustained a long-term contact of igneous minerals with liquid water, altering such igneous minerals into clays²⁵. This process could either have been maintained at the surface, if the climate was warm enough, or have occurred through the actions of fluids in a warm, shallow crust. In both cases, martian clays are likely to record an alteration having taken place over geologic timescales with liquid water present at thermodynamic equilibrium. However, the formation of clays on Mars by impact and volcanic hydrothermal activity has been discussed by several authors^{31,32}. Such processes, which do not require liquid water to be stable at the surface, could account for the hydrated silicates identified at least in some specific areas.

Finally, it is important to note that most phyllosilicate deposits are distinct from sulphate deposits as mapped by OMEGA: in general, they do not occur together (in a few cases, such as a few deposits within Aram Chaos, Terra Meridiani and the Becquerel crater, sulphates and clays might be mixed). In contrast to clays, sulphate formation is favoured under acidic water conditions, and does not necessarily imply the long-term presence of liquid water¹⁰. The two major families of alteration products detected by OMEGA—phyllosilicates and sulphates—could thus trace two different processes separated in time, referring to two major climatic episodes in the history of Mars: an early Noachian Mars, resulting in the formation of hydrated silicates, followed by a more acidic environment in which sulphates formed, rather than clays.

Received 28 July; accepted 27 September 2005.

- Gendrin, A. *et al.* Sulfates in Martian layered terrains: the OMEGA/Mars Express view. *Science* **307**, 1587–1591 (2005).
- Langevin, Y., Poulet, F., Bibring, J.-P. & Gondet, B. Sulfates in the North polar region of Mars detected by OMEGA/Mars Express. *Science* **307**, 1584–1586 (2005).
- Squyres, S. *et al.* The Opportunity rover's Athena Science Investigation at Meridiani Planum, Mars. *Science* **306**, 1698–1703 (2004).
- Toulmin, P. *et al.* Geochemical and mineralogical interpretation of the Viking inorganic chemical results. *J. Geophys. Res.* **82**, 4625–4634 (1977).
- Bridges, J. C. *et al.* Alteration assemblages in martian meteorites: implications for near-surface processes. *Space Sci. Rev.* **96**, 365–392 (2001).
- Clark, R. N., Swayze, G. A., Singer, R. B. & Pollack, J. B. High-resolution reflectance spectra of Mars in the 2.3-micron region—Evidence for the mineral scapolite. *J. Geophys. Res.* **95**, 14463–14480 (1990).
- Wyatt, M. B. & McSweeney, H. Y. Spectral evidence for weathered basalt as an alternative to andesite in the northern lowlands of Mars. *Nature* **417**, 263–266 (2002).
- Bandfield, J. L. Global mineral distributions on Mars. *J. Geophys. Res.* **107**, doi:10.1029/2001JE001510 (2002).
- Kraft, M. D., Michalski, J. R. & Sharp, T. G. Effects of pure silica coatings on thermal emission spectra of basaltic rocks: Considerations for Martian surface mineralogy. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL018848 (2003).
- Bibring, J.-P. *et al.* Mars surface diversity as revealed by the OMEGA/Mars Express observations. *Science* **307**, 1576–1581 (2005).
- Mustard, J. *et al.* Olivine and pyroxene diversity in the crust of Mars. *Science* **307**, 1594–1597 (2005).
- Bishop, J. L. & Pieters, C. M. Low-temperature and low atmospheric pressure infrared reflectance spectroscopy of Mars soil analog materials. *J. Geophys. Res.* **100**, 5369–5379 (1995).
- Bishop, J. L., Murad, E. & Dyar, M. D. The influence of octahedral and tetrahedral cation substitution on the structure of smectites and serpentines as observed through infrared spectroscopy. *Clay Mineral.* **37**, 617–628 (2002).
- Bishop, J. L., Madejová, J., Komadel, P. & Fröschl, H. The influence of structural Fe, Al, and Mg on the infrared OH bands in spectra of dioctahedral smectites. *Clay Mineral.* **37**, 607–616 (2002).
- Clark, R. N., King, T. V. V., Klejwa, M., Swayze, G. A. & Vergo, N. High spectral resolution reflectance spectroscopy of minerals. *J. Geophys. Res.* **95**, 12653–12680 (1990).
- Keeling, J., Mauger, A. & Raven, M. in *Regolith 2004* (ed. Roach, I. C.) 166–170 (Cooperative Research Centre/Landscape Environments and Mineral Exploration (CRC/LEME), 2004).
- Frost, R. L., Klopogge, J. T. & Ding, Z. Near-infrared spectroscopy study of nontronites and ferruginous smectite. *Spectrochim. Acta A* **58**, 1657–1668 (2002).
- Hiesinger, H. & Head, J. W. The Syrtis Major volcanic province, Mars: Synthesis from Mars Global Surveyor data. *J. Geophys. Res.* **109**, E01004, doi:10.1029/2003JE002143 (2004).
- Hoefen, T. M. *et al.* Discovery of olivine in the Nili Fossae region of Mars. *Science* **302**, 627–630 (2003).
- Christensen, P. R. *et al.* Morphology and composition of the surface of Mars: Mars Odyssey THEMIS results. *Science* **300**, 2056–2061 (2003).
- Hamilton, V. E. & Christensen, P. R. Evidence for extensive, olivine-rich bedrock on Mars. *Geology* (in the press) (2005).
- Malin, M. C. & Edgett, K. S. Sedimentary rocks of early Mars. *Science* **290**, 1927–1937 (2000).
- Edgett, K. S. & Parker, T. J. Water on early Mars: Possible subaqueous sedimentary deposits covering ancient cratered terrain in western Arabia and Sinus Meridiani. *Geophys. Res. Lett.* **24**, 2897–2900 (1997).
- Ivanov, M. A. & Head, J. W. Chryse Planitia, Mars: Topographic configuration, outflow channel continuity and sequence, and tests for hypothesized ancient bodies of water using Mars Orbiter Laser Altimeter (MOLA) data. *J. Geophys. Res.* **106**, 3275–3295 (2001).
- Velde, B., Righi, D., Meunier, A., Hillier, S. & Inoue, A. in *Origin and Mineralogy of Clays* (ed. Velde, B.) 8–42 (Springer, Berlin, 1995).
- Righi, D. & Meunier, A. in *Origin and Mineralogy of Clays* (ed. Velde, B.) 43–157 (Springer, Berlin, 1995).
- Schiffman, P., Spero, H. J., Southard, R. J. & Swanson, D. A. Controls on palagonitization versus pedogenic weathering of basaltic tephra: Evidence from the consolidation and geochemistry of the Keanakakoi Ash Member, Kilauea Volcano. *Geochem. Geophys. Geosyst.* **1**, doi:10.1029/2000GC000068 (2000).
- Gooding, J. L. Chemical weathering on Mars—Thermodynamic stabilities of primary minerals and their alteration products from mafic igneous rocks. *Icarus* **33**, 483–513 (1978).
- Allen, P. A. *Earth Surface Processes* (Blackwell Science, Malden, Massachusetts, 1997).
- Dickinson, W. W. & Rosen, M. R. Antarctic permafrost: an analogue for water and diagenetic minerals on Mars. *Geology* **31**, 199–202 (2003).
- Newsom, H. E. Hydrothermal alteration of impact melt sheets with implications for Mars. *Icarus* **44**, 207–216 (1980).
- Rathbun, J. A. & Squyres, S. W. Hydrothermal systems associated with Martian impact craters. *Icarus* **157**, 362–372 (2002).

Acknowledgements Mars Express is operated by ESA/ESOC. Laboratory reflectance data used in this paper are provided through cooperation between the OMEGA team and NASA and the Mars Reconnaissance Orbiter (MRO) Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) team. Reflectance measurements were acquired at the NASA Reflectance Experiment Laboratory (RELAB) at Brown University.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.P. (francois.poulet@ias.fr).

The Omega Team M. Berthé¹, J.-P. Bibring¹, Y. Langevin¹, S. Erard¹, O. Forni¹, A. Gendrin¹, B. Gondet¹, N. Manaud¹, F. Poulet¹, G. Poulléau¹, A. Soufflot¹, M. Combes², P. Drossart², T. Encrenaz², T. Fouchet², R. Melchiorri², G. Bellucci³, F. Altieri³, V. Formisano³, S. Fonti³, F. Capaccioni³, P. Cerroni³, A. Coradini³, O. Korabev⁴, V. Kottsov⁴, N. Ignatiev⁴, D. Titov⁴, L. Zasova⁴, N. Mangold⁵, P. Pinet⁶, B. Schmitt⁷, C. Sotin⁸, E. Hauber⁹, H. Hoffmann⁹, R. Jaumann⁹, U. Keller¹⁰, R. Arvidson¹¹, J. Mustard¹² & F. Forget¹³

Affiliations for participants: ¹IAS, Université Paris-Sud, F-91405, France. ²LESIA, Observatoire de Paris/Meudon, 92195 Meudon, France. ³IFSI-INAF, Via del Fosso del Cavaliere, 00133 Rome, Italy. ⁴Space Research Institute (IKI) 117997, 84/32 Profsoyuznaya Str, Moscow, Russia. ⁵IDES, Université Paris-Sud, F-91405 Orsay, France. ⁶Observatoire Midi-Pyrénées, 31000 Toulouse, France. ⁷Laboratoire de Planétologie, 38400 Toulouse, France. ⁸Planétologie, Université de Nantes, 44322 Nantes, France. ⁹DLR, Berlin, 12489, Germany. ¹⁰MPAE, Lindau, Germany. ¹¹Earth and Planetary Sciences, Washington University, Saint-Louis, Missouri 63130, USA. ¹²Geological Sciences, Brown University, Providence, Rhode Island 02912, USA. ¹³LMD, Université de Paris 6, 75252, Paris, France.

ARTICLES

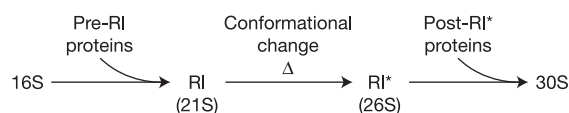
An assembly landscape for the 30S ribosomal subunit

Megan W. T. Talkington^{1†}, Gary Siuzdak¹ & James R. Williamson¹

Self-assembling macromolecular machines drive fundamental cellular processes, including transcription, messenger RNA processing, translation, DNA replication and cellular transport. The ribosome, which carries out protein synthesis, is one such machine, and the 30S subunit of the bacterial ribosome is the preeminent model system for biophysical analysis of large RNA–protein complexes. Our understanding of 30S assembly is incomplete, owing to the challenges of monitoring the association of many components simultaneously. Here we have developed a method involving pulse–chase monitored by quantitative mass spectrometry (PC/QMS) to follow the assembly of the 20 ribosomal proteins with 16S ribosomal RNA during formation of the functional particle. These data represent a detailed and quantitative kinetic characterization of the assembly of a large multicomponent macromolecular complex. By measuring the protein binding rates at a range of temperatures, we find that local transformations throughout the assembling subunit have similar but distinct activation energies. Thus, the prevailing view of 30S assembly as a pathway proceeding through a global rate-limiting conformational change must give way to one in which the assembly of the complex traverses a landscape dotted with various local conformational transitions.

The assembly of the 30S ribosomal subunit is a complex dance of macromolecular folding and binding in which 20 proteins bind to rRNA as it folds, creating a complete particle^{1–3} that is competent to participate in translation of mRNA. Assembly *in vitro* has shown that secondary structure in the 16S rRNA (local helices) is stabilized by Mg²⁺-containing buffer alone, but tertiary (long-range) folding depends on the proteins⁴. Because protein binding sites are created as the rRNA folds, ribosomal protein binding reports on local rRNA tertiary conformation throughout assembly^{5–9}. Much knowledge of the order and mechanism of 30S assembly has thus been amassed by identifying the proteins bound at equilibrium in incomplete assembly reactions^{10,11}.

A slow rate-limiting folding transition has long been inferred from the observation that incomplete particles with an altered sedimentation coefficient (21 S versus 30 S) form at low temperatures (0–15 °C; refs 12–14). Heating these incomplete particles, termed the reconstitution intermediate (RI), shifts their sedimentation coefficient to 26 S (RI*) and enables them to complete assembly at low temperatures. The RI → RI* transition, which was thought to be a conformational change in the rRNA, was proposed to be the rate-limiting step of assembly even at higher temperatures, because the apparent concentration independence of the overall assembly rate suggested a unimolecular rate-limiting step¹². The RI → RI* transition characterizes the canonical scheme of 30S assembly, which has remained essentially unchanged for 35 years:



The next step in characterizing the mechanism of 30S assembly is to determine the kinetics by which the various proteins bind to the assembling subunit. Because standard methods are not capable of directly monitoring the binding of many proteins simultaneously,

however, we have developed a method, PC/QMS (pulse–chase monitored by quantitative mass spectrometry), that measures the kinetics of binding the individual proteins during assembly of the whole complex.

A method for studying assembly of the whole 30S subunit

PC/QMS takes advantage of the ability of mass spectrometry to quantify large numbers of proteins relative to stable isotope-labelled species, an approach that is widely used in proteomics^{15–18}. Assembly of 30S subunits is initiated by incubating the *Escherichia coli* 16S rRNA of 1,542 nucleotides with a mixture of uniformly ¹⁵N-labelled 30S proteins (S2–S21)¹⁹. At various time points, binding of the ¹⁵N-proteins is chased with an excess of unlabelled (¹⁴N) proteins. Completely formed 30S subunits are purified, and the ¹⁵N/¹⁴N ratio for each protein is determined by matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry (MALDI-TOF^{20–22}; Fig. 1a). The ¹⁵N/¹⁴N ratios can be quantified accurately, as judged by standard curves collected on known mixtures of labelled and unlabelled proteins (Fig. 1b), and most of the 30S proteins are observed in a single scan (Fig. 1c). The assay has been validated by measuring the binding rate of the *Aquifex aeolicus* S15 protein to a 16S rRNA fragment using PC/QMS and comparing the results with those of a gel mobility shift assay (see Supplementary Information and Fig. S1). Plotting the fractional isotope ratios for a given protein as a function of time produces a progress curve for the binding of that protein during assembly of the whole subunit. In this way, the binding kinetics of all of the ribosomal proteins can be determined in a single experiment.

Protein binding rates match the existing 30S assembly map

Under standard conditions (see Methods), similar to those identified as optimal for *in vitro* assembly¹², the proteins bind with rates distributed throughout two orders of magnitude (Fig. 2a–c). The trends in these data correspond well to protein binding rates inferred

¹Departments of Molecular Biology and Chemistry, and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, USA. †Present address: Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA.

from the reactivity of 16S rRNA nucleotides to chemical probes over time⁶ and to the binding order determined in classical equilibrium experiments¹⁰ (Fig. 2b). Assembly *in vitro* maintains the 5' to 3' directionality and overall protein binding order, including late assembly of the interdomain junction that forms the site of mRNA decoding (Fig. 2c), that is observed *in vivo*^{6,23}, despite taking place on a mature 16S rRNA rather than on a nascent precursor rRNA transcript^{24,25}.

Folding and binding occur at similar rates

Characterizing the mechanism of 30S assembly requires measurement of the protein binding kinetics under various conditions, and PC/QMS is sufficiently rapid to permit the collection of data sets in several different conditions. To begin probing the nature of the rate-limiting steps of assembly, we varied the concentration of rRNA and proteins in the assembly reaction. At one extreme, if binding is the rate-limiting step for a particular protein, then the binding rate should be directly proportional to the concentration. At the other extreme, if a unimolecular folding event is rate-limiting, then the rate should be insensitive to concentration. In fact, the intermediate situation is observed for many proteins—that is, the protein binding rates are weakly affected by concentration (Fig. 3)—which indicates that RNA folding and protein binding occur at similar rates. All of the

proteins observed here show some concentration dependence in their binding rates; thus, folding does not seem to be rate-limiting for any of them.

30S assembly proceeds through many rate-limiting transitions

To characterize assembly intermediates, we measured protein binding rates at low temperature, where RI has been found to accumulate. The ¹⁵N-protein pulse was done at low temperature, and the temperature was restored to the optimum (40 °C) upon addition of the ¹⁴N-protein chase. Consistent with previous measurements of overall assembly rates¹², protein binding is slow at 15 °C (Fig. 4a), requiring more than 2 days to proceed to completion. Unexpectedly, none of the proteins is disproportionately slowed as compared with the others and none plateaus at a low extent of binding—an observation that initially seems to be inconsistent with stalling of assembly at a 21S intermediate (RI).

The standard RI → RI* mechanism, whereby assembly stalls at the 21S intermediate at low temperatures, implies that the late proteins have much lower rates of binding than the early proteins at low temperatures, whereas the binding rates for all proteins are more similar at 40 °C, where assembly proceeds smoothly. The temperature dependence of the protein binding rates is characterized by the Arrhenius activation energy (E_a), and there are generally two ways to explain the previous observations in terms of activation energies. Either the activation energies are much larger for the late binding proteins than for the early binding proteins, or there is a change in the rate-determining step for the late proteins to a process with a larger activation energy at low temperatures.

The temperature dependence of the binding rate of each protein was measured over the accessible range (Fig. 4b), and the activation energies were determined from the slopes of the Arrhenius plots (Fig. 4b, c). The activation energies are generally similar for all of the proteins, being scattered throughout a relatively narrow range of ~24–44 kcal mol⁻¹. The binding activation energies observed are all similar to the E_a for overall assembly of 38 kcal mol⁻¹ determined previously (ref. 12). The magnitude of the activation energies corresponds to the melting of about four RNA base pairs²⁶ and also corresponds to the activation energy for the folding of small proteins, although we cannot at present determine the relative contributions of RNA and protein folding to the kinetics observed. Although there is a rough trend that the activation energies are slightly larger for the late binding proteins than for the early binding proteins, the correlation is poor, and the magnitude of the differences in activation energy is insufficient to produce stalling of assembly at low temperature. Furthermore, the Arrhenius plots are linear over the accessible range (see Methods), which clearly indicates that the activation energies do not change with temperature and thus that the rate-determining step is the same for each protein at high and low temperature.

Consequently, no one step is solely responsible for the apparent E_a of overall assembly. The slowly binding proteins, which include both those that precede the canonical RI → RI* transition and those that follow it, do not have the highest E_a values (Fig. 4c), so the last steps of assembly are not more dependent on temperature than the earlier steps. Furthermore, the rates and E_a values of the slowly binding proteins are not well correlated, indicating that the final stages of assembly are limited by many different transitions. Until now, there has been no way to follow these different transitions because the individual protein binding rates have not been determined during assembly at several temperatures. PC/QMS has enabled us to do this, and we find that the classic RI → RI* mechanism is not adequate to explain the rates and activation energies observed for binding of the individual proteins.

These observations suggest that although a 21S 'particle' can be isolated from assembly at low temperature, this 21S particle is not a true assembly 'intermediate'. It seems likely that the reason that 21S particles are retrieved from sucrose gradient purification of

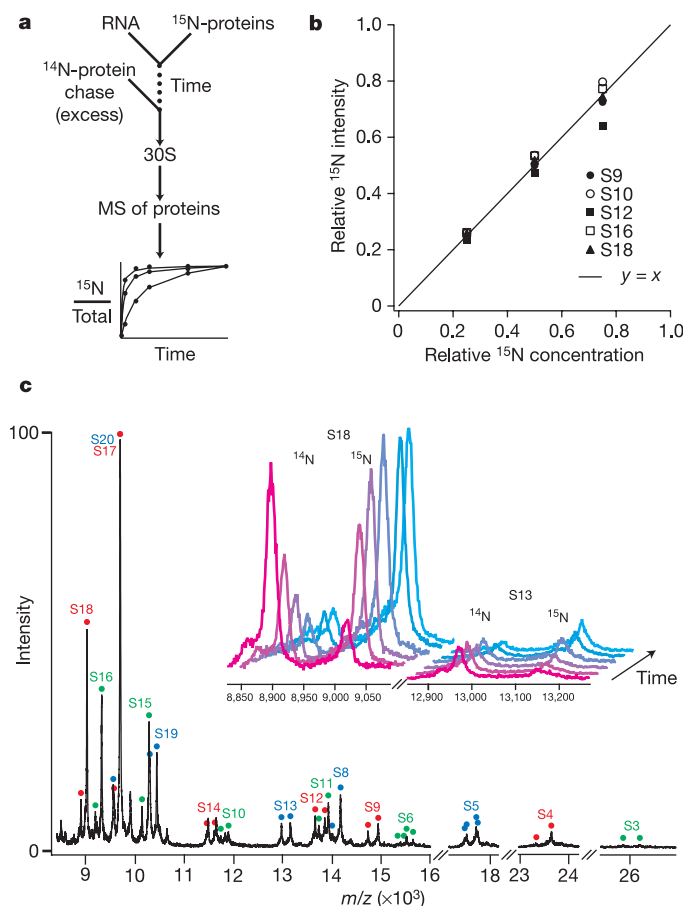


Figure 1 | The PC/QMS method for measuring protein binding kinetics in the 30S ribosomal subunit. **a**, The PC/QMS method. **b**, Quantification of relative ¹⁵N-protein concentrations for several proteins from three standard mixtures of native ¹⁵N- and ¹⁴N-30S subunits. The average relative intensities for all proteins from the three mixtures were 0.24 ± 0.03 , 0.50 ± 0.03 and 0.73 ± 0.04 (mean $\pm$ s.d.). **c**, MALDI-TOF mass spectrum of 30S proteins from the 2-min time point of an assembly reaction done under standard conditions. Inset shows expanded spectra for several time points for proteins S18 and S13. Additional details are provided in the Supplementary Information.

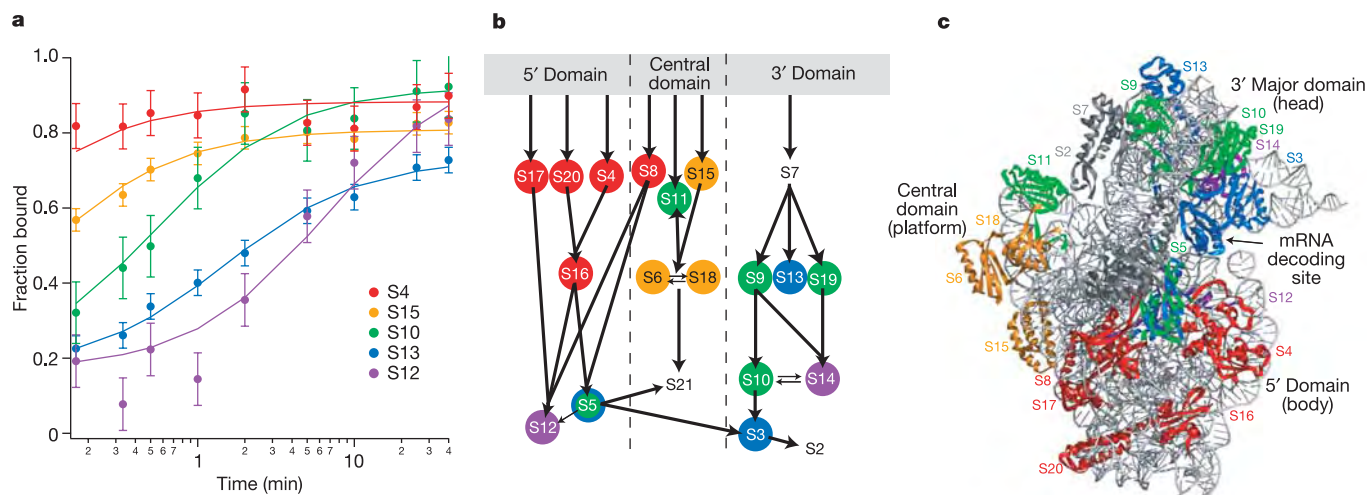


Figure 2 | Binding kinetics of 30S proteins measured using PC/QMS under standard conditions. **a**, Representative progress curves for protein binding (see Supplementary Fig. S2), fitted as described in the Methods. The error bars are derived from the s.d. of standard samples (see Supplementary Information). **b**, Proteins in the Nomura assembly map (refs 10, 11; and S. C. Agalarov & J.R.W., unpublished data) are coloured by their binding

rates (see Supplementary Table S1). Red, 20 to $\geq 30 \text{ min}^{-1}$; orange, 8.1–15 min^{-1} ; green, 1.2–2.2 min^{-1} ; blue, 0.38–0.73 min^{-1} ; purple, 0.18–0.26 min^{-1} . S5 is shown in green and blue to represent the binding rates of the unacetylated and acetylated forms, respectively. Grey bar represents 16S rRNA. **c**, Proteins in an X-ray crystal structure of the 30S subunit from *T. thermophilus*¹, coloured as in **b**.

low-temperature assembly reactions is that a diverse collection of unstable particles that are in the process of assembling all sediment at $\sim 21 \text{ S}$ until they accomplish a transition that shifts them to 26 S . This depiction agrees with the previous observations that the characteristics of RI are variable and that some pre-RI proteins bind only transiently at the RI stage¹³. It is likely that weakly bound proteins dissociate to different extents during the PC/QMS chase as compared with sucrose gradient centrifugation, such that the binding of some

‘pre-RI’ proteins (particularly S5, S12 and S19) is observed to be slow by PC/QMS.

The slight clustering in protein binding rates at 15°C (Fig. 4a, c) may indicate the presence of populated assembly intermediates. Because the members of a group do not share the same activation energy (Fig. 4c), however, it seems that the binding of the proteins in a given group are not all limited by a single RNA folding step. Assembly by various local transitions rather than a single, global step enables the various subunits in a population to assemble into the native structure by various routes rather than a requisite pathway. Equilibrium footprinting of reconstituted RI and RI* particles indicates that conformational changes are scattered throughout the 16S rRNA sequence, although centred on the active site¹⁴. This observation is consistent with the presence of many local conformational changes that may take place in parallel during late stages of assembly. Thus, just as macromolecular folding pathways have been expanded to folding landscapes that can be traversed by any of various parallel pathways^{27–30}, so too can the assembly of a multi-component complex, the 30S subunit, now be represented by a landscape (Fig. 5).

An assembly landscape for the 30S subunit

In the landscape representation all possible conformations of the 16S rRNA map onto a free-energy surface, but in the absence of proteins the native 30S conformation is energetically unfavourable. Folding can proceed along many possible pathways to the native state because the landscape is composed of many local and modest barriers. A unique feature of the 30S landscape, as compared with unimolecular folding landscapes²⁷, is the intermolecular protein binding, which alters the shape of the free-energy surface during the assembly process. Once RNA folding produces a new binding site, protein binding creates new downhill directions by which further RNA folding can proceed. The marked alteration of the 16S folding landscape that accompanies ribosomal protein binding is analogous to the changes in protein folding landscapes that occur on shifting from denaturing to native conditions. Each protein binding event further stabilizes the native 30S conformation until all assembly pathways converge at this state. Despite the changes in the landscape that accompany protein binding, the heights of the various barriers encountered on any particular pathway seem to be similar.

Viewing 30S assembly as a landscape is supported not only by the

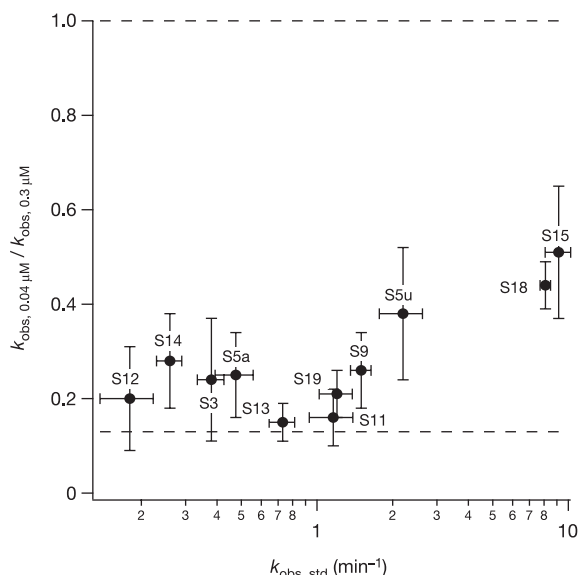


Figure 3 | Ratio of the protein binding rates observed at two concentrations versus the rates at standard concentration. Ratios of 1.0 or 0.13 (broken lines) would indicate unimolecular or bimolecular rate-limiting steps, respectively. The errors in k_{obs} (s.d. from the fits of progress curves) are propagated to produce the errors bars. Proteins that bind very rapidly at the standard concentration are not shown, because the rates cannot be accurately determined from the present data. S10 data are not shown owing to a poor signal. Proteins S6 and S8 have high ratios, similar to two other central domain proteins, S18 and S15. Proteins S16, S17 and S20 have lower ratios, similar to most proteins.

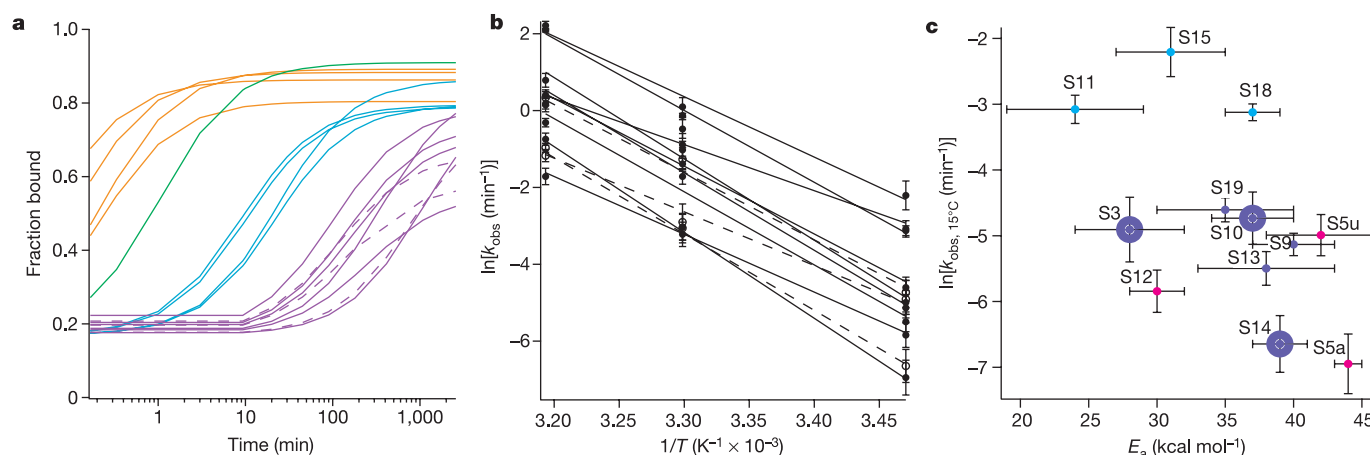


Figure 4 | The temperature dependence of protein binding rates. **a**, Fits of binding progress curves at 15 °C, coloured according to the rates (see Supplementary Table S1). Orange, 4.4–21 min^{-1} ; green, 1.0 min^{-1} ; cyan, 0.044–0.11 min^{-1} ; purple, 0.00096–0.010 min^{-1} . Post-RI* proteins (S3, S10 and S14) are shown as broken lines here and in **b**. **b**, Arrhenius plots of the observed rates (see Supplementary Fig. S3). The errors in k_{obs} (s.d. from the

fits of progress curves) are propagated to produce the error bars. Proteins that bind very rapidly are not shown here or in **c**. **c**, Protein binding rates at 15 °C versus the activation energies (see Supplementary Table S1). The errors in E_a are the s.d. from the linear Arrhenius plot fits. Proteins are coloured according to the 30S domain (magenta, 5' domain; cyan, central domain; purple, 3' domain). Post-RI* proteins have large points.

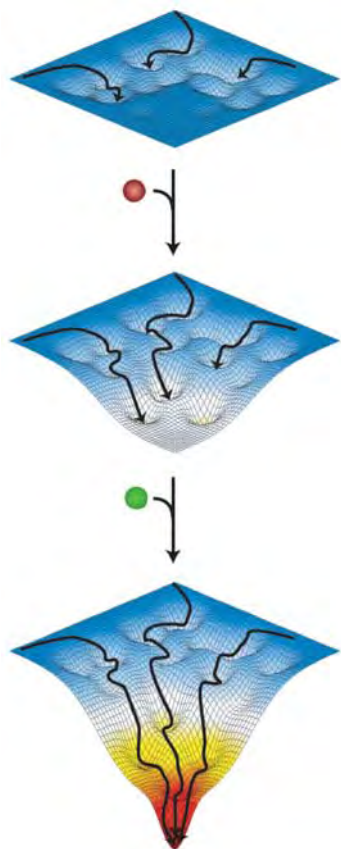


Figure 5 | An assembly landscape for 30S assembly. The horizontal axes of the surface correspond to 16S rRNA conformational space, and the vertical axis is free energy. The native conformation of the 16S rRNA adopted in the 30S subunit is located at the bottom corner. In the absence of proteins, this is not the lowest-energy conformation of the RNA. Parallel folding pathways are indicated by the arrows on the energy surface. Local folding creates protein binding sites, and large changes in the landscape accompany protein binding (coloured spheres). Sequential protein binding eventually stabilizes the native 30S conformation. All pathways converge on this point, and there is no bottleneck through which all folding trajectories must pass.

detailed kinetic data reported here, but also by the classical equilibrium data summarized in the assembly map (Fig. 2b), which show that the ribosomal proteins do not have an absolute dependence on each other for binding, but rather can bind in various orders³¹. Indeed, Nomura and colleagues predicted that assembly actually proceeds by several pathways even as they proposed the simple RI → RI* model, because it was observed that different proteins potentiated the formation of RI* particles to varying extents¹³.

Assembly by a global rate-limiting step, which would be represented by a bottleneck on the landscape, could bring assembly to a standstill under non-optimal conditions. Assembly through a landscape of different barriers, by contrast, would mean that slowing any one of the steps would slow down, but not completely stall, assembly. Such a robust assembly landscape is surely one of many functions encoded by strongly conserved ribosomal sequences. RNA and protein chaperones are expected to have a role in assembly, and the protein chaperone DnaK has been specifically implicated in aiding 30S assembly^{32–34}. The landscape model developed here predicts that there are many folding transitions that are points at which chaperones might assist.

The assay introduced here, PC/QMS, has made it possible to begin to construct an assembly landscape for a large macromolecular complex, the 30S ribosomal subunit. The assay reports the kinetics at which different sites throughout the 30S subunit assemble, and it can be conducted under various conditions designed to mimic the intracellular assembly reaction, as well as with 30S components engineered to assess the roles of particular components and functional groups. We expect that themes from the 30S assembly landscape will inform our understanding of the assembly of ribonucleoproteins and of large complexes in general. As a general method suitable for studying site-specific assembly in multicomponent complexes, PC/QMS can also be adapted to these systems.

METHODS

Pulse-chase assembly of 30S subunits. Mixtures of all 30S proteins (unlabelled or ^{15}N -labelled) and 16S rRNA were prepared from native 30S subunits (see Supplementary Information). Binding titrations indicated that the concentrations of active proteins in the mixtures were roughly stoichiometric (within ~2-fold); thus, differences in the concentrations of the proteins should have a minimal effect on the binding rates observed. The standard assembly conditions were 0.3 μM 16S and 0.45 μM ^{15}N -proteins in assembly buffer containing 25 mM Tris-HCl (pH 7.5 at room temperature), 330 mM KCl, 20 mM MgCl_2 and 2 mM dithiothreitol at 40 °C (ref. 35); the chase was $5 \times$ unlabelled proteins. Non-specific binding of the excess proteins in the chase was resolved by purifying the

assembled 30S subunits in 10–40% sucrose gradients containing a high salt concentration (assembly buffer plus 0.5 M NH_4Cl). Particles assembled under standard conditions in the presence of excess proteins and purified in high-salt conditions are properly formed, as judged by the extent to which they bind 50S subunits to form particles that migrate as 70S particles³⁶. (Assembled subunits are slightly less active than native 30S; assembled and chased subunits are as active as those that are not chased.)

The PC/QMS assay was done at 40, 30 and 15 °C. The very low rate of assembly at low temperatures makes 15 °C the lowest temperature at which it is practical to measure binding kinetics; over the course of a 6-d experiment at 10 °C, some precipitation was observed in protein samples, causing concerns about the integrity of samples over the long periods of time required for assembly at such low temperatures (see Supplementary Table S1).

MALDI analysis. The proteins bound during the pulse-chase reaction were extracted from the assembled 30S subunits (see Supplementary Information). The extracted proteins were analysed with a Voyager-DE STR MALDI-TOF mass spectrometer (PerSeptive Biosystems) operated in linear mode. The intensities of the protein peaks were determined by fitting each peak to a single gaussian function using Igor Pro (WaveMetrics). The heights of the gaussian fits (after background subtraction; see Supplementary Information) were taken as the peak intensities. We report the relative ^{15}N -labelled protein intensities: ^{15}N -protein/(^{14}N -protein + ^{15}N -protein).

Analysis of protein binding progress curves. The progress curves of relative ^{15}N -protein intensity versus time were fitted to an equation of two-state binding for a bimolecular system, $\text{R} + \text{P} \rightarrow \text{RP}$:

$$\text{d}[\text{RP}]/\text{d}t = k_{\text{on}}[\text{R}][\text{P}] - k_{\text{off}}[\text{RP}]$$

where R is 16S rRNA, P is one of the proteins, and RP is the complex. Because ^{15}N -protein binding was chased with 5×10^{-4} M ^{14}N -proteins, the minimum fraction of ^{15}N -protein bound was 0.17 ($1/(1 + 5) = 0.17$). The binding rate observed is the product of k_{on} and the total RNA concentration ($k_{\text{obs}} = k_{\text{on}}[\text{R}]$). For most proteins, this observed binding rate probably represents many rate constants—binding of the protein itself as well as earlier proteins and rRNA folding.

Arrhenius analysis. The activation energies of protein binding are calculated from the slopes of the Arrhenius plots using the Arrhenius equation, $k = \text{Ae}^{-E_a/RT}$.

Received 14 July; accepted 22 September 2005.

- Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).
- Schluenzen, F. *et al.* Structure of functionally activated small ribosomal subunit at 3.3 Å resolution. *Cell* **102**, 615–623 (2000).
- Brodersen, D. E., Clemons, W. M. Jr, Carter, A. P., Wimberly, B. T. & Ramakrishnan, V. Crystal structure of the 30S ribosomal subunit from *Thermophilus*: structure of the proteins and their interactions with 16S rRNA. *J. Mol. Biol.* **316**, 725–768 (2002).
- Moazed, D., Stern, S. & Noller, H. F. Rapid chemical probing of conformation in 16S ribosomal RNA and 30S ribosomal subunits using primer extension. *J. Mol. Biol.* **187**, 399–416 (1986).
- Agalarov, S. C. & Williamson, J. R. A hierarchy of RNA subdomains in assembly of the central domain of the 30S ribosomal subunit. *RNA* **6**, 402–408 (2000).
- Powers, T., Daubresse, G. & Noller, H. F. Dynamics of *in vitro* assembly of 16S rRNA into 30S ribosomal subunits. *J. Mol. Biol.* **232**, 362–374 (1993).
- Recht, M. I. & Williamson, J. R. Thermodynamics and kinetics of central domain assembly. *Cold Spring Harbor Symp. Quant. Biol.* **66**, 591–598 (2001).
- Stagg, S. M., Mears, J. A. & Harvey, S. C. A structural model for the assembly of the 30S subunit of the ribosome. *J. Mol. Biol.* **328**, 49–61 (2003).
- Stern, S., Powers, T., Changchien, L. M. & Noller, H. F. RNA–protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. *Science* **244**, 783–790 (1989).
- Held, W. A., Ballou, B., Mizushima, S. & Nomura, M. Assembly mapping of 30S ribosomal proteins from *Escherichia coli*. *J. Biol. Chem.* **249**, 3103–3111 (1974).
- Culver, G. M. Assembly of the 30S ribosomal subunit. *Biopolymers* **68**, 234–249 (2003).
- Traub, P. & Nomura, M. Structure and function of *Escherichia coli* ribosomes VI. Mechanism of assembly of 30S ribosomes studied *in vitro*. *J. Mol. Biol.* **40**, 391–413 (1969).
- Held, W. A. & Nomura, M. Rate-determining step in the reconstitution of *Escherichia coli* 30S ribosomal subunits. *Biochemistry* **12**, 3273–3281 (1973).
- Holmes, K. L. & Culver, G. M. Mapping structural differences between 30S ribosomal subunit assembly intermediates. *Nature Struct. Mol. Biol.* **11**, 179–186 (2004).
- Gygi, S. P. *et al.* Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnol.* **17**, 994–999 (1999).
- Oda, Y., Huang, K., Cross, F. R., Cowburn, D. & Chait, B. T. Accurate quantitation of protein expression and site-specific phosphorylation. *Proc. Natl Acad. Sci. USA* **96**, 6591–6596 (1999).
- Pasa-Tolic, L. *et al.* High throughput proteome-wide precision measurements of protein expression using mass spectrometry. *J. Am. Chem. Soc.* **121**, 7949–7950 (1999).
- Sechi, S. & Oda, Y. Quantitative proteomics using mass spectrometry. *Curr. Opin. Chem. Biol.* **7**, 70–77 (2003).
- Traub, P. & Nomura, M. Structure and function of *E. coli* ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. *Proc. Natl Acad. Sci. USA* **59**, 777–784 (1968).
- Beavis, R. C. & Chait, B. T. Matrix-assisted laser desorption/ionization mass-spectrometry of proteins. *Methods Enzymol.* **270**, 519–551 (1996).
- Hillenkamp, F. & Karas, M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods Enzymol.* **193**, 280–295 (1990).
- Jennings, K. R. & Dolnikowski, G. G. Mass analyzers. *Methods Enzymol.* **193**, 37–61 (1990).
- Pichon, J., Marvaldi, J. & Marchis-Mouren, G. The *in vivo* order of protein addition in the course of *Escherichia coli* 30S and 50S subunit biogenesis. *J. Mol. Biol.* **96**, 125–137 (1975).
- Mangiarotti, G., Apirion, D., Schlessinger, D. & Silengo, L. Biosynthetic precursors of 30S and 50S ribosomal particles in *Escherichia coli*. *Biochemistry* **7**, 456–472 (1968).
- Srivastava, A. K. & Schlessinger, D. in *The Ribosome: Structure, Function, & Evolution* (ed. Hill, W. *et al.*) 426–434 (American Society of Microbiology, Washington DC, 1990).
- Xia, T. *et al.* Thermodynamic parameters for an expanded nearest-neighbour model for formation of RNA duplexes with Watson–Crick base pairs. *Biochemistry* **37**, 14719–14735 (1998).
- Dill, K. A. & Chan, H. S. From Levinthal to pathways to funnels. *Nature Struct. Biol.* **4**, 10–19 (1997).
- Pan, J., Thirumalai, D. & Woodson, S. A. Folding of RNA involves parallel pathways. *J. Mol. Biol.* **273**, 7–13 (1997).
- Rook, M. S., Treiber, D. K. & Williamson, J. R. Fast folding mutants of the *Tetrahymena* group I ribozyme reveal a rugged folding energy landscape. *J. Mol. Biol.* **281**, 609–620 (1998).
- Thirumalai, D. & Woodson, S. A. Kinetics of folding of proteins and RNA. *Acc. Chem. Res.* **29**, 433–439 (1996).
- Nomura, M. Assembly of bacterial ribosomes. *Science* **179**, 864–873 (1973).
- Alix, J. H. & Guerin, M. F. Mutant DnaK chaperones cause ribosome assembly defects in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **90**, 9725–9729 (1993).
- Maki, J. A., Schnobrich, D. J. & Culver, G. M. The DnaK chaperone system facilitates 30S ribosomal subunit assembly. *Mol. Cell* **10**, 129–138 (2002).
- Maki, J. A., Southworth, D. R. & Culver, G. M. Demonstration of the role of the DnaK chaperone system in assembly of 30S ribosomal subunits using a purified *in vitro* system. *RNA* **9**, 1418–1421 (2003).
- Serdnyuk, I. N., Agalarov, S. C., Sedelnikova, S. E., Spirin, A. S. & May, R. P. Shape and compactness of the isolated ribosomal 16S RNA and its complexes with ribosomal proteins. *J. Mol. Biol.* **169**, 409–425 (1983).
- Culver, G. M. & Noller, H. F. Efficient reconstitution of functional *Escherichia coli* 30S ribosomal subunits from a complete set of recombinant small subunit ribosomal proteins. *RNA* **5**, 832–843 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the staff of the TSRI Center for Mass Spectrometry for assistance with mass spectrometry; M. I. Recht, S. C. Agalarov and S. P. Ryder for discussions and technical assistance; the laboratories of D. B. Goodin, S. P. Mayfield and A. Schneemann for use of equipment; and M. J. Fedor, J. D. Puglisi and S. P. Ryder for critically reading the manuscript. This work was supported by a grant from the NIH (to J.R.W.) and by predoctoral fellowships from the NSF and the Skaggs Institute for Chemical Biology (to M.W.T.T.).

Author Contributions The experimental work in this manuscript was carried out by M.W.T.T., with advice and support from G.S. and J.R.W.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.R.W. (jrwil@scripps.edu).

Lipid-protein interactions in double-layered two-dimensional AQP0 crystals

Tamir Gonen¹, Yifan Cheng¹, Piotr Sliz^{2,3}, Yoko Hiroaki⁴, Yoshinori Fujiyoshi⁴, Stephen C. Harrison^{2,3} & Thomas Walz¹

Lens-specific aquaporin-0 (AQP0) functions as a specific water pore and forms the thin junctions between fibre cells. Here we describe a 1.9 Å resolution structure of junctional AQP0, determined by electron crystallography of double-layered two-dimensional crystals. Comparison of junctional and non-junctional AQP0 structures shows that junction formation depends on a conformational switch in an extracellular loop, which may result from cleavage of the cytoplasmic amino and carboxy termini. In the centre of the water pathway, the closed pore in junctional AQP0 retains only three water molecules, which are too widely spaced to form hydrogen bonds with each other. Packing interactions between AQP0 tetramers in the crystalline array are mediated by lipid molecules, which assume preferred conformations. We were therefore able to build an atomic model for the lipid bilayer surrounding the AQP0 tetramers, and we describe lipid-protein interactions.

Members of the aquaporin (AQP) family form membrane pores that are either highly selective for water (aquaporins) or also permeable to other small neutral solutes such as glycerol and urea (aquaglyceroporins) (reviewed in ref. 1). Structural studies have revealed that all AQPs share the same basic architecture, which consists of two tandem repeats, each containing a bundle of three transmembrane α -helices and a hydrophobic loop with the highly conserved asparagine-proline-alanine (NPA) motif^{2–8}. The two NPA-containing loops B and E fold back into the membrane and form short α -helices (HB and HE) that line the water pore. The ar/R constriction site, so named because it is formed by an aromatic residue and an arginine residue, confers water selectivity to AQP pores, whereas the NPA motifs are important in the proton exclusion mechanism (reviewed in ref. 9).

AQP0 is the most abundant protein in lens fibre cell membranes, where it forms not only water pores but also the 11–13-nm ‘thin lens junctions’ that assemble after proteolytic cleavage of the cytoplasmic termini^{10,11}. We recently presented the structure of the AQP0-mediated membrane junction at 3 Å resolution as determined by electron crystallography of double-layered two-dimensional (2D) crystals⁷. The structure showed that AQP0 junctions are stabilized by specific interactions between tetramers in adjoining membranes involving proline residues almost exclusively. Calculated pore profiles also showed that the pore in junctional AQP0 is highly constricted by a substantially extended ar/R constriction site and a novel second constriction site⁷, which may be involved in the pH regulation of AQP0 water conductance¹².

The 1.9 Å structure of junctional AQP0

The water pore in junctional AQP0 seen at 3 Å resolution appears closed⁷. However, the water molecules were not resolved and we were unable to demonstrate directly the absence of water molecules from the AQP0 pore. Using a better batch of 2D crystals (Fig. 1a), the carbon sandwich specimen preparation technique¹³ and a helium-cooled top-entry 300-kV electron microscope¹⁴, equipped with a field

emission gun and a 4,096-pixel $\times$ 4,096-pixel charge-coupled device camera, we have now been able to collect electron diffraction data to much higher resolution.

Electron diffraction patterns collected from untilted specimens (Fig. 1b) as well as highly tilted specimens (Supplementary Fig. 1) showed strong and sharp diffraction spots to a resolution beyond 2 Å. The final data set comprised 286 diffraction patterns recorded at tilt angles of up to 71.3° that were merged to a resolution of 1.7 Å. The structure was refined with CNS¹⁵ to a resolution of 1.9 Å using scattering factors for 300 kV electrons. The final refinement statistics are summarized in Table 1. At the improved resolution of 1.9 Å, water molecules can be identified clearly in the density map, and the rings of many aromatic residues are represented by doughnut-shaped densities (Fig. 1c).

Junctional versus non-junctional AQP0

The formation of thin junctions in the lens core correlates with proteolytic cleavage of at least a fraction of the AQP0 (refs 11, 16, 17). A comparison of the structure of AQP0 in the 2D crystals we describe here with that of AQP0 in three-dimensional (3D) crystals, as reported⁸, allows us to identify differences between junctional and non-junctional conformations. In the 3D crystals, uncleaved AQP0 tetramers do not form junctions, and the pores are filled with water molecules⁸. In the double-layered 2D crystals, which form only from the partly proteolytically cleaved AQP0 population purified from lens core, the tetramers interact with each other through their extracellular surfaces (Supplementary Fig. 2b) and thus recapitulate the arrangement in the thin junctions between lens fibre cells¹¹. In the junctional conformation the water pores are closed.

When we inspected the 3D crystal structure (Protein Data Bank accession code 1YMG) with PROCHECK¹⁸ and WHATIF¹⁹, we found that the unit-cell dimensions required adjustment (from $a = b = 110.53$ Å and $c = 53.39$ Å to $a = b = 109.53$ Å and $c = 52.82$ Å). We refined the structure into the adjusted cell and noticed continuous density for both termini (not built in the

¹Department of Cell Biology and ²Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA. ³Howard Hughes Medical Institute and Children's Hospital Laboratory of Molecular Medicine, 320 Longwood Avenue, Boston, Massachusetts 02115, USA.

⁴Department of Biophysics, Kyoto University, Oiwake, Kitashirakawa Sakyo-ku, Kyoto 606-8502, Japan.

deposited model). We therefore modelled N-terminal residues 2–5 and C-terminal residues 240–263. The C terminus mediates important crystal packing contacts with two neighbouring tetramers (Supplementary Fig. 2a). Both termini also engage in interactions that seem to be of functional importance, as described in more detail below. The refinement (see Supplementary Table 1) improved the G factor¹⁸ from -1.51 to 0.21 .

In intact, non-junctional AQP0, the N-terminal and C-terminal regions have ordered conformations (Fig. 2b, c). Arg 226 and Lys 228 near the C terminus interact electrostatically with Ser 79, Gln 80 and Asp 150; the N terminus loops back and tucks Trp 2 into a hydrophobic pocket lined by Phe 9, Trp 10 and Leu 84 (Fig. 2c). The resulting conformation allows Glu 3 to interact with Ser 240, thereby bridging between the N-terminal and C-terminal segments. All these interactions are eliminated in truncated AQP0 (Fig. 2d, e). Cleavage at residue 234 causes the remainder of the C-terminal segment to move away from the membrane surface, disrupting the previous contacts of Arg 226 and Lys 228 (Fig. 2d). The N-terminal and C-terminal cleavages also eliminate Glu 3 and Ser 240.

How do these changes at the cytoplasmic face affect junctional interactions at the extracellular face? Disruption of the network of interactions involving the two termini seems to correlate with rearrangements in extracellular loop A (Fig. 2a). Pro 38 is particularly critical. In the non-junctional structure, this residue points away from the centre of the tetramer (Fig. 2f); moreover, Trp 34 lies above the pore and projects outward, blocking the approach of a second tetramer, and Arg 33 intervenes between two monomers. In the truncated, junctional AQP0 tetramer (Fig. 2g), loop A has reconfigured, positioning Pro 38 so that it can form a rosette-like structure at the centre of the tetramer and mediate a major junctional contact. Arg 33 and Trp 34 also swap positions, so that Trp 34 no longer interferes with the close approach of another tetramer. In the completed junction, all three residues interact with the corresponding residues from the apposing tetramer (the two Arg residues interact through a water molecule) (Supplementary Fig. 3).

Water molecules in the AQP0 pore

The water pore in non-junctional AQP0 contains seven water molecules (Fig. 3a, left); that in junctional AQP0 contains only three (Fig. 3a, right). The two pores have the same diameter over much of their lengths, but the pore in junctional AQP0 is narrower at the positions of the two constriction sites (Fig. 3a,

centre). Constriction site I (CS-I) in non-junctional AQP0 spans 3 Å and has a minimum diameter of 2.31 Å; CS-I of junctional AQP0 extends over 10 Å and the pore narrows to 1.33 Å. The large difference in the length of CS-I is due mainly to the side chain of Met 176, which extends into the pore in junctional AQP0 (Supplementary Fig. 4a) but points away from it in non-junctional AQP0, allowing access by additional water molecules (Supplementary Fig. 4b). Constriction site II (CS-II) is narrower in junctional AQP0 (diameter 1.37 Å) than in non-junctional AQP0 (diameter 1.75 Å). The constricted pore in junctional AQP0 can thus accommodate only three water molecules, which seem to be trapped in the closed pore because it narrows above and below them (Supplementary Fig. 4c, d).

In our initial report of the closed water pore in junctional AQP0 (ref. 7), we proposed that AQP0 and other aquaporins might be in a dynamic equilibrium between an open and a closed pore conformation. We also suggested that pore closure might be triggered by the stabilization of an alternative conformation of Arg 187 (part of the ar/R constriction site) seen in the structure of junctional AQP0. A recent molecular dynamics study supports this notion: it showed that Arg 189 in AQPZ (corresponding to Arg 187 in AQP0) could adopt two conformations²⁰. The 'UP' state, which is seen in most AQP crystal structures, had an open pore, filled by a continuous single file of water. The 'DOWN' state, seen in our structure of junctional AQP0, had a pore completely blocked by the Arg side chain, and prolonged blockage resulted in the loss of all water molecules from the pore. Although attractive, a conformational switch of the arginine in the ar/R constriction site cannot be the only mechanism for AQP gating, because Arg 187 is in the 'DOWN' state not only in our closed, junctional AQP0 but also in the open, non-junctional AQP0 structure⁸. The main difference between the open pore and the closed pore lies in the conformation of the side chain of Met 176 (see above), a residue not present in AQPZ.

The distances between the three water molecules (4 Å or more) in the closed pore are too long for hydrogen bonding (Fig. 3b, right, and Supplementary Fig. 5, right). The water coordinated to the Asn residues of the two NPA motifs donates a hydrogen bond to the hydroxyl group of Tyr 24, which in turn donates a hydrogen bond to the water molecule in the extracellular half of the water pathway (Fig. 3b, right, and Supplementary Fig. 5, right). The corresponding two water molecules in the open water pore of non-junctional AQP0 have the same hydrogen-bonding pattern (Fig. 3b, left, and Supplementary Fig. 5, left), and all the other water molecules are within

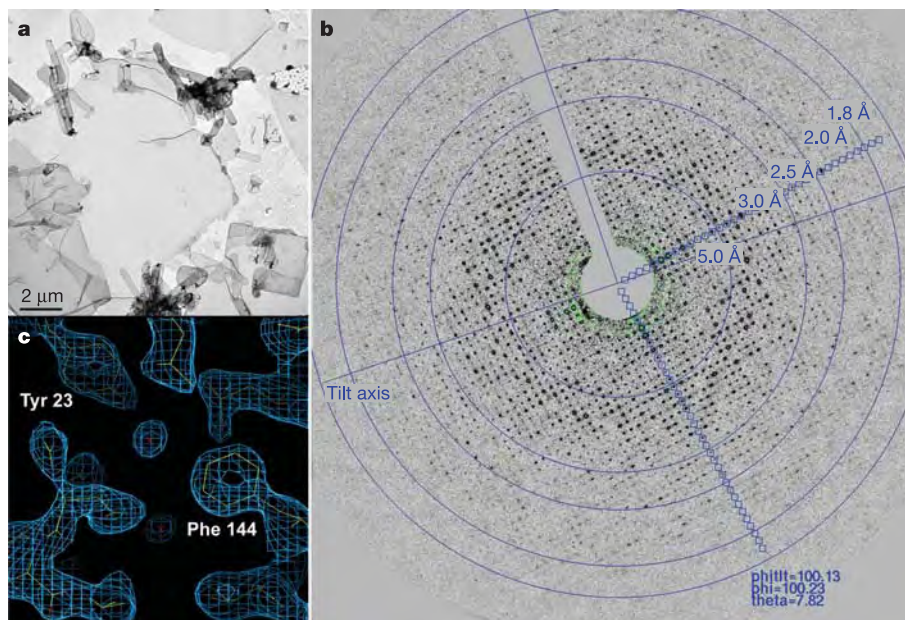


Figure 1 | Electron crystallography of AQP0 junctions. **a**, Double-layered AQP0 2D crystals were often several micrometres in size. **b**, A typical electron diffraction pattern recorded from an untilted AQP0 2D crystal prepared by the carbon sandwich technique¹³, showing diffraction spots to a resolution beyond 2 Å. **c**, Region of the final $2F_o - F_c$ map of AQP0 refined to 1.9 Å resolution. Two aromatic residues, Tyr 23 and Phe 144, that line the water pore in AQP0 are represented by doughnut-shaped densities.

Table 1 | Electron crystallographic data

Two-dimensional crystals	
Layer group	<i>p</i> 422
Unit cell (Å)	<i>a</i> = <i>b</i> = 65.5
Thickness (assumed) (Å)	160
Electron diffraction	
Number of patterns merged	286 (0°, 11; 20°, 43; 45°, 107; 60°, 87; 70°, 38)
Resolution limit for merging (Å)	1.7
<i>R</i> _{Friedel} (%)	14.25
<i>R</i> _{merge} (%)	16.60
Observed amplitudes to 1.9 Å	126,980
Unique reflections	22,293
Maximum tilt angle (°)	71.3
Fourier space sampled	80.0% (70.5% at 2.0–1.9 Å)
Multiplicity	5.7 (2.5 at 2.0–1.9 Å)
Crystallographic refinement (5.0–1.9 Å)	
Resolution limit for refinement (Å)	1.9
Crystallographic <i>R</i> factor (%)	25.81
Free <i>R</i> factor (%)	29.93
Reflections in working/test set	14,600/1,580
Non-hydrogen protein atoms	1,784
Non-hydrogen lipid atoms	348
Solvent molecules	76
Average protein <i>B</i> factor (Å ²)	48.4
Ramachandran plot (%)	97.5; 2.5; 0 (allowed; generous; disallowed)

*R*_{free} is calculated from a randomly chosen 10% of reflections, and *R*_{cryst} is calculated over the remaining 90% of reflections.

hydrogen-bonding distance of each other. This 'phenolic barrier' created by Tyr 24, a residue not seen in the other known AQP structures, may be responsible for the poor water conductance of AQP0 in comparison with other AQPs, which contain a continuous line of hydrogen-bonded water molecules. The space occupied by Tyr 24 might also explain why the open AQP0 pore contains only

seven water molecules, whereas molecular dynamics studies showed eight water molecules in AQP1 (refs 21, 22) and AQPZ²⁰ and nine in GlpF²³.

AQP0 water conductance is dependent on pH, with a maximum at pH 6.5 and only about half the activity at pH 10.5 (ref. 12). These conductance characteristics are not changed by proteolytic cleavage of AQP0 (ref. 24). As our structure, obtained with the double-layered 2D crystals grown at pH 6 (ref. 7), reveals fewer water molecules in the pore than the structure determined from the 3D crystals grown at pH 10.5 (ref. 8), pore closure seems to be a result of junction formation, not pH shift.

Lipid-protein interactions

Crystals of membrane proteins occasionally contain lipid molecules. For example, the structure of bacteriorhodopsin from lipid cubic-phase crystallization revealed 13 phytanyl lipids, 7 of which formed a bilayer structure, and a squalene²⁵. These, and all other lipids found in crystal structures so far (78 lipids in total²⁶), originate from the native membrane, from which they co-purify with the crystallized membrane protein. Such tightly bound lipids have been found to be essential for the structural integrity and activity of a number of membrane proteins²⁷.

None of the AQP 3D crystals examined so far contain lipids, and 2D crystals of AQPs can form with a variety of different lipids, indicating that AQPs might have neither a requirement for specific lipids nor high-affinity lipid-binding sites. Nevertheless, our density map revealed that between the AQP0 tetramers are horseshoe-shaped features characteristic of lipid molecules (Fig. 4a). Indeed, close inspection revealed that lipids bridge all the contacts between tetramers within a layer and that the tetramers have essentially no direct lateral interaction. In composite omit maps we could identify nine lipids per AQP0 monomer, which we modelled as complete or

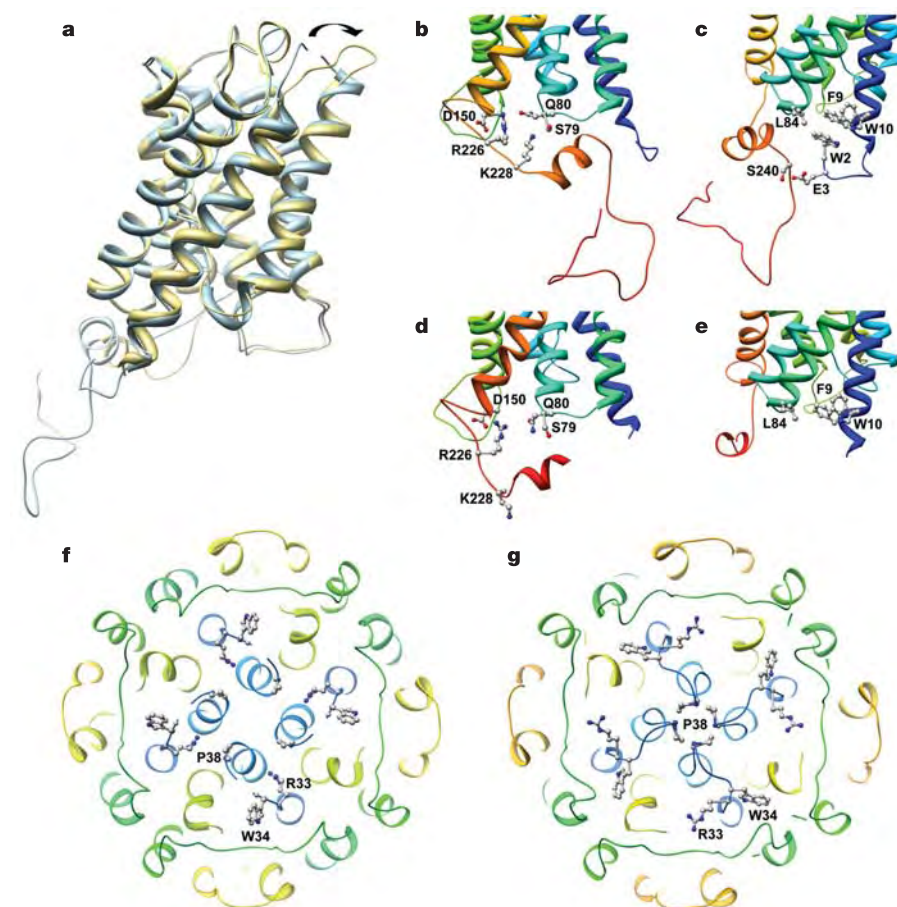


Figure 2 | Structural differences between junctional and non-junctional AQP0. **a**, X-ray structure of non-junctional AQP0 (blue) superimposed on the electron-microscope structure of junctional AQP0 (yellow). The arrow indicates the conformational switch of extracellular loop A. **b**, Interactions involving the C terminus of full-length AQP0. **c**, Interactions involving the N terminus of full-length AQP0. **d**, C terminus of cleaved AQP0. **e**, N terminus of cleaved AQP0. **f**, View of the extracellular surface of the X-ray structure of non-junctional, full-length AQP0. **g**, View of the extracellular surface of the electron-microscope structure of junctional AQP0 from the lens core, showing the rosette-like structure formed by the Pro 38 residues.

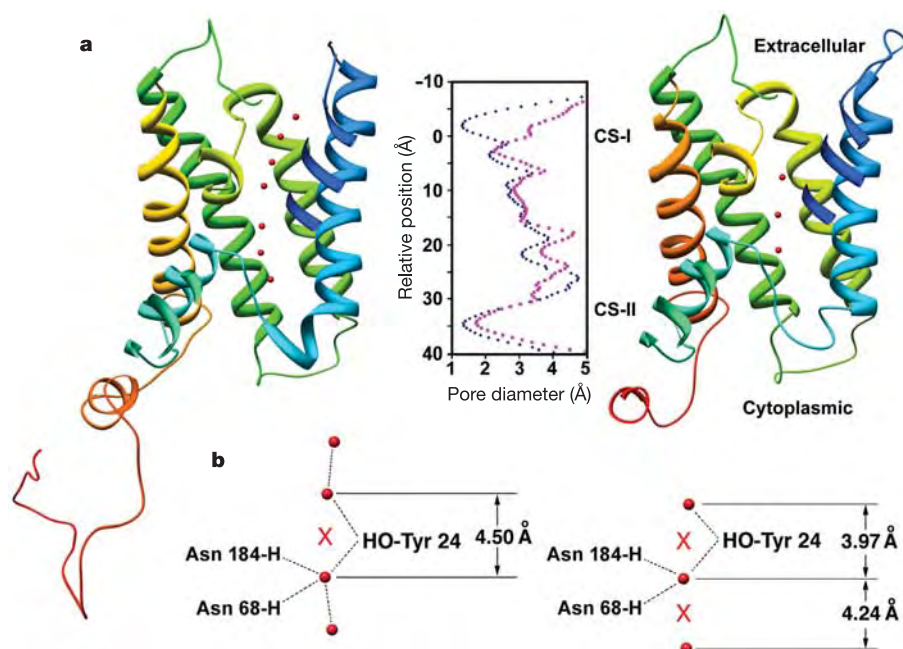


Figure 3 | The water pore in AQP0. **a**, The pore in non-junctional AQP0 (left) contains seven water molecules (red spheres), whereas the pore in junctional AQP0 contains only three water molecules (right). Calculated pore profiles (middle) confirm that the pore in junctional AQP0 (purple) is more constricted than that in non-junctional AQP0 (pink). **b**, Hydrogen-

bonding pattern of water molecules in the pore of non-junctional AQP0 (left) and junctional AQP0 (right). The hydrogen-bonding network is disrupted by Tyr 24, which introduces a phenolic barrier. In junctional AQP0 all three water molecules are too far apart to form hydrogen bonds. Dotted lines represent hydrogen bonds. See Supplementary Fig. 5.

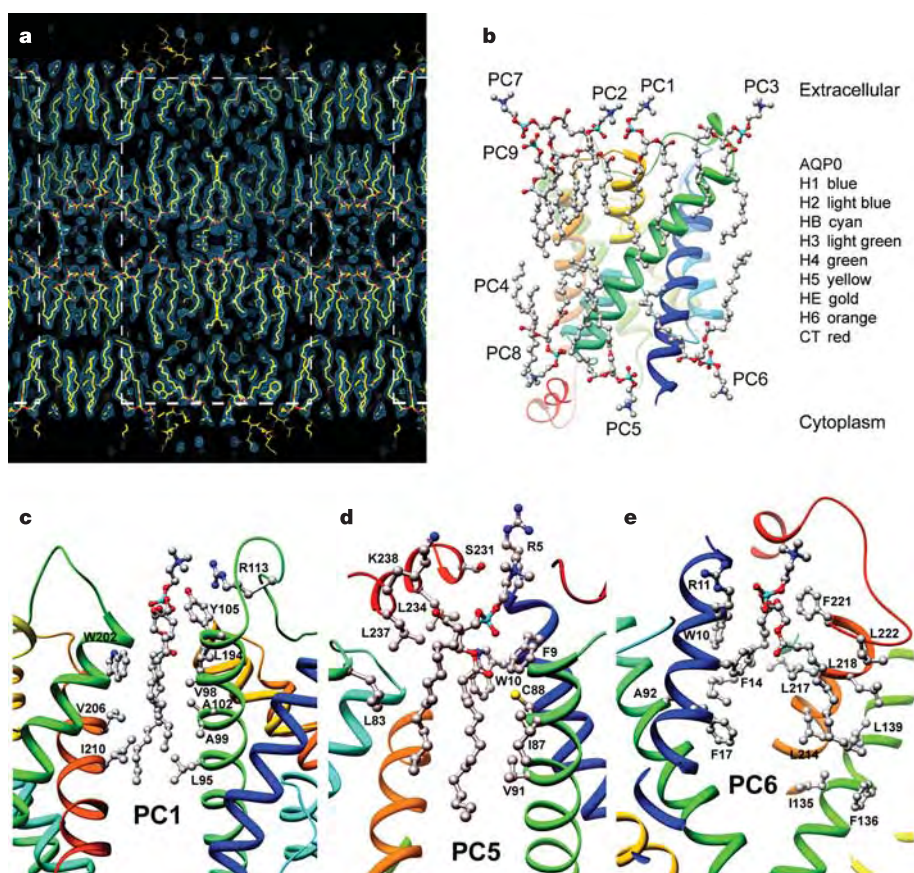


Figure 4 | Lipid-protein interactions in double-layered AQP0 2D crystals. **a**, Vertical slab through the $2F_o - F_c$ density map with modelled lipid molecules, revealing the two lipid bilayers in the double-layered AQP0 2D crystal. **b**, The nine lipids surrounding an AQP0 monomer in the 2D crystal. Lipids PC1 to PC7 are annular lipids, whereas lipids PC8 and PC9 are bulk

lipids with no direct protein contacts. H1–H6, helices 1–6; CT, cytoplasmic tail. See Supplementary Fig. 6 for a stereo view. **c–e**, Three examples of lipids sandwiched between two AQP0 molecules. The acyl chains of PC1 adopt a closed conformation (**c**), those of PC5 a slightly splayed conformation (**d**) and those of PC6 a widely splayed conformation (**e**).

partial molecules of dimyristoyl phosphatidyl choline (DMPC, the lipid used for 2D crystallization) (Fig. 4a). Phospholipid headgroups have a chiral centre at C2 of the glycerol, and the DMPC we used is a racemic mixture. Density is weak or absent at most C2 positions in our map, and often at the attached ester group as well, indicating that there might be little or no selectivity for the biological enantiomer. Very strong density for the phosphate groups, weaker but well-defined density for the trimethylamine groups of the cholines, and unambiguous density for the acyl chains allowed us to build and refine a model in which we chose an enantiomer for each lipid more or less arbitrarily. We have not yet attempted to refine the two alternatives with 50% occupancy each. We have annotated these lipids as PC1 to PC9 (Fig. 4b and Supplementary Fig. 6). PC1 to PC7 have extensive protein contacts and seem to represent 'annular lipids' immediately adjacent to a membrane-embedded protein. PC8 and PC9 are not in contact with protein and thus represent bulk lipids. A detailed description of lipid-protein contacts is provided in Supplementary Materials. As AQP0 has no tight lipid-binding sites, interactions between the annular lipids and the AQP0 subunits are likely to represent the kinds of contact that occur between any membrane protein and the lipids surrounding it.

Annular lipids must adapt to the irregular surface of a transmembrane protein to create a smooth interface for bulk lipids. This fit limits the mobility (and perhaps the chemistry) of annular lipids, because their conformations are partly defined by the protein surface. In our 2D arrays, most of the annular lipids are sandwiched between two tetramers and thus mediate lattice interactions (Supplementary Fig. 7). This packing further restricts their conformations. The cell dimensions of our reconstituted junctions are the same as those in thin junctions between lens fibre cells²⁸. We therefore suggest that the lipid-protein interactions we observe in our 2D crystals with the artificial lipid DMPC are representative of those formed by AQP0 tetramers with native lipids in lens fibre cell membranes.

The lipids form a one-molecule-wide annular shell around the protein. The positions of the headgroups vary by only ± 2 Å in the direction perpendicular to the membrane plane, with a separation of about 34 Å from phosphate to phosphate. The dimensions of the bilayer correspond closely to those of fully hydrated, fluid-phase DMPC²⁹. A hydrated network of hydrogen bonds and salt bridges holds the lipid phosphates in place. Protein groups interacting with phosphates include three arginine side chains, a tyrosine hydroxyl group that mediates one of the arginine contacts, a lysine, a tryptophan indole nitrogen, a glutamine side-chain amide, and at least one main-chain amide. Similar interactions have been described for specifically bound lipids³⁰.

Acyl chains fill the gaps between adjacent tetramers. Their conformations clearly adapt to the knobs and grooves of the apposed hydrophobic protein surfaces. Figure 4c–e illustrates three examples. PC1 in the extracellular leaflet is the best ordered of the nine DMPC molecules. Its acyl chains are nearly fully extended, packed against those of PC2 and PC3 and sandwiched between five nonpolar side chains from one AQP0 molecule and three from the other. PC5 in the cytoplasmic leaflet has somewhat less extended acyl chains. The phosphate receives a hydrogen bond from the indole nitrogen of Trp 10 and Lys 238 (as well as the poorly ordered N-terminal segment) of an adjacent subunit. The acyl chains, packed between those of PC4 and PC6, contact four hydrophobic side chains from one subunit (including the hydrophobic face of Trp 10) and three from another. PC6, also in the cytoplasmic leaflet, has widely splayed acyl chains, separated by side chains from the two apposed AQP0 molecules. Phe 14 of one molecule and Leu 217 of another are in van der Waals contact through the gap: this is the only direct interaction between tetramers within a layer.

PC8 and PC9 lie near the four-fold axis. They do not contact protein and thus represent bulk lipids. Neither is as well ordered as the annular lipids. Indeed, PC8 (in the cytoplasmic leaflet) is probably only statistically ordered (two, rather than four, molecules

about a four-fold axis), because there is space for only one of the two acyl chains and no density for the headgroup. The headgroup of PC9 lies about 3 Å closer to the midplane of the bilayer than those of the four other extracellular leaflet lipids; the bilayer thickness may therefore be influenced by adjacency to the protein.

Discussion

AQP0 serves a dual function in the lens, acting both as a water channel and as an adhesion molecule. During the differentiation of fibre cells, and as they grow older and become buried more deeply in the lens, AQP0 is cleaved at both termini^{16,17,31,32}. This processing seems to be the trigger for junction formation^{10,11}. Comparison of our structure of junctional AQP0 with that of non-junctional AQP0 (ref. 8) indicates that cleavage of the two cytoplasmic termini might translate into a conformational switch in extracellular loop A, eliminating steric hindrance from Trp 34 and allowing Pro 38 to stabilize the junctional interaction. Formation of the junction also seems to correlate with changes in the side-chain positions of residues lining the pore, most importantly in Met 176, resulting in substantial constriction. Three water molecules are trapped in the centre of the water pathway, too far apart from each other to be linked by hydrogen bonds.

Mutations in AQP0 result in cataracts^{33–37}. Most of these mutations interfere with the correct trafficking of AQP0 to the plasma membrane (reviewed in ref. 38), rather than with efficient water conduction as proposed recently⁸. However, the mutations might prevent proper interaction with lipid, which in many instances has a key role in the folding and integration of membrane proteins into a bilayer^{30,39,40}.

The lipids in our 2D crystals indeed demonstrate a well-defined role for annular lipids in forming a boundary for the bilayer-inserted protein. The headgroup interactions of these lipids resemble those described for specifically bound lipids³⁰, and the acyl chains are tightly packed around the laterally projecting hydrophobic side chains of the protein. We suggest that when junctions form between lens fibre cells, the annular lipids already partly immobilized by interaction with AQP0 mediate the lattice contacts, just as DMPC does in our reconstituted junctions. Although the composition of natural membrane lipids is far less homogeneous than in our crystals, the incorporation of both DMPC enantiomers shows that the headgroup interactions are somewhat adjustable, and imperfect crystallinity of the acyl chains indicates that C₁₆ or C₁₈ chains, or even unsaturated chains, could readily be accommodated. It remains to be determined whether AQP0 exhibits selectivity for its annular lipids and hence for the lipid composition within the junctional lattices.

METHODS

Purification and 2D crystallization of AQP0 from the lens core. AQP0 was purified from the core of sheep lenses and reconstituted into 2D crystals as described previously⁷. The 2D crystals were grown at a lipid:protein ratio of 0.25 (by mass), which corresponds to a molar ratio of 37 lipids per AQP0 tetramer. This is very close to the number of lipids, 36, that we could model per AQP0 tetramer.

Electron microscopy and data processing. Negatively stained samples were prepared and imaged as described previously⁷. Specimens for cryoelectron microscopy were prepared as described¹³. In brief, double-layered AQP0 2D crystals were mixed with an equal amount of 10% trehalose and the suspension was applied to a molybdenum grid covered with a thin carbon film. The grid was blotted to remove excess material, and a second carbon film was placed on top of the sample. Grids were plunged into liquid nitrogen and loaded into a JEM3000SFF electron microscope equipped with a top-entry helium stage and operated at an acceleration voltage of 300 kV (ref. 14). Low-dose electron diffraction patterns were recorded with a 4,096-pixel $\times$ 4,096-pixel charge-coupled-device camera (Gatan). Electron diffraction patterns were analysed and merged as described⁴¹.

Molecular replacement, model building and refinement. The structure of AQP0 was determined by molecular replacement as described previously⁷. Crowther fast cross-rotation function calculations identified an orientation of a single subunit in the asymmetric unit as the top solution with a rotation

function signal double the value of the second peak. The translation function gave a top solution with an *R* factor of 42.7% and a correlation coefficient of 84.3%. The model was refined with CNS version 1.1 (ref. 15). Rigid-body refinement was used to evaluate what scattering factors to use. In a resolution range of 10–2.5 Å, both X-ray and electron scattering factors for 120-kV electrons⁴¹ produced comparable results. In a resolution range of 10–1.9 Å, electron scattering factors for 300-kV electrons produced the best result (*R* = 35.92%), in comparison with electron scattering factors for 120-kV electrons (*R* = 67.06%) and X-ray scattering factors (*R* = 37.39%). For all subsequent refinement steps, refinement was performed in a resolution range of 5–1.9 Å without bulk solvent flattening, using a unit cell of *a* = *b* = 65.5 Å, an assumed thickness of 160 Å, and electron scattering factors for 300-kV electrons. Model building was performed with O⁴² using $2F_o - F_c$ density maps and simulated annealing composite omit maps. Protein residues 5–239 were visible and were modelled (major cleavage sites *in vivo*³²), together with nine lipid (DMPC) and 76 water molecules. The model was refined by cycles of simulated annealing, *B*-factor refinement, and remodelling. The final refinement statistics are presented in Table 1.

For the crystal structure of non-junctional AQP0 (ref. 8) we downloaded the structure factors and coordinates from the Protein Data Bank (accession code 1YMG). The structure was evaluated with PROCHECK¹⁸ and WHATIF¹⁹ and the unit cell dimensions were adjusted from *a* = *b* = 110.531 Å and *c* = 53.390 Å to *a* = *b* = 109.531 Å and *c* = 52.822 Å. A composite omit map was calculated and a clear continuous density was observed for both the N terminus and the C terminus. We modelled N-terminal residues 2–5 and C-terminal residues 240–263 in O⁴² and refined the structure in CNS version 1.1 (ref. 15) by cycles of simulated annealing, *B*-factor refinement, and remodelling. The final refinement statistics are presented in Supplementary Table 1.

Received 6 September; accepted 12 October 2005.

- Agre, P. *et al.* Aquaporin water channels—from atomic structure to clinical medicine. *J. Physiol. (Lond.)* **542**, 3–16 (2002).
- Murata, K. *et al.* Structural determinants of water permeation through aquaporin-1. *Nature* **407**, 599–605 (2000).
- Fu, D. *et al.* Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481–486 (2000).
- Sui, H., Han, B. G., Lee, J. K., Walian, P. & Jap, B. K. Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**, 872–878 (2001).
- Ren, G., Reddy, V. S., Cheng, A., Melnyk, P. & Mitra, A. K. Visualization of a water-selective pore by electron crystallography in vitreous ice. *Proc. Natl Acad. Sci. USA* **98**, 1398–1403 (2001).
- Savage, D. F., Egea, P. F., Robles-Colmenares, Y., O'Connell, J. D. & Stroud, R. M. Architecture and selectivity in aquaporins: 2.5 Å X-ray structure of aquaporin Z. *PLoS Biol.* **1**, E72 (2003).
- Gonen, T., Sliz, P., Kistler, J., Cheng, Y. & Walz, T. Aquaporin-O membrane junctions reveal the structure of a closed water pore. *Nature* **429**, 193–197 (2004).
- Harries, W. E., Akhavan, D., Miercke, L. J., Khademi, S. & Stroud, R. M. The channel architecture of aquaporin O at a 2.2-Å resolution. *Proc. Natl Acad. Sci. USA* **101**, 14045–14050 (2004).
- de Groot, B. L. & Grubmüller, H. The dynamics and energetics of water permeation and proton exclusion in aquaporins. *Curr. Opin. Struct. Biol.* **15**, 176–183 (2005).
- Kistler, J. & Bullivant, S. Lens gap junctions and orthogonal arrays are unrelated. *FEBS Lett.* **111**, 73–78 (1980).
- Gonen, T., Cheng, Y., Kistler, J. & Walz, T. Aquaporin-O membrane junctions form upon proteolytic cleavage. *J. Mol. Biol.* **342**, 1337–1345 (2004).
- Nemeth-Cahalan, K. L., Kalman, K. & Hall, J. E. Molecular basis of pH and Ca²⁺ regulation of aquaporin water permeability. *J. Gen. Physiol.* **123**, 573–580 (2004).
- Gyobu, N. *et al.* Improved specimen preparation for cryo-electron microscopy using a symmetric carbon sandwich technique. *J. Struct. Biol.* **146**, 325–333 (2004).
- Fujiyoshi, Y. The structural study of membrane proteins by electron crystallography. *Adv. Biophys.* **35**, 25–80 (1998).
- Brunker, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
- Roy, D., Spector, A. & Farnsworth, P. N. Human lens membrane: comparison of major intrinsic polypeptides from young and old lenses isolated by a new methodology. *Exp. Eye Res.* **28**, 353–358 (1979).
- Takemoto, L., Takehana, M. & Horwitz, J. Covalent changes in MIP26K during aging of the human lens membrane. *Invest. Ophthalmol. Vis. Sci.* **27**, 443–446 (1986).
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
- Vriend, G. WHAT IF: a molecular modeling and drug design program. *J. Mol. Graph.* **8**, 526–529 (1990).
- Wang, Y., Schulten, K. & Tajkhorshid, E. What makes an aquaporin a glycerol channel? A comparative study of AqpZ and GlpF. *Structure* **13**, 1107–1118 (2005).
- de Groot, B. L. & Grubmüller, H. Water permeation across biological membranes: mechanism and dynamics of aquaporin-1 and GlpF. *Science* **294**, 2353–2357 (2001).
- Zhu, F., Tajkhorshid, E. & Schulten, K. Molecular dynamics study of aquaporin-1 water channel in a lipid bilayer. *FEBS Lett.* **504**, 212–218 (2001).
- Tajkhorshid, E. *et al.* Control of the selectivity of the aquaporin water channel family by global orientational tuning. *Science* **296**, 525–530 (2002).
- Ball, L. E. *et al.* Water permeability of C-terminally truncated aquaporin O (AQP0 1–243) observed in the aging human lens. *Invest. Ophthalmol. Vis. Sci.* **44**, 4820–4828 (2003).
- Luecke, H., Schobert, B., Richter, H. T., Cartailler, J. P. & Lanyi, J. K. Structure of bacteriorhodopsin at 1.55 Å resolution. *J. Mol. Biol.* **291**, 899–911 (1999).
- Wiener, M. in *Protein-Lipid Interactions: From Membrane Domains to Cellular Networks* (ed. Tamm, L. K.) 29–49 (Wiley-VCH, Weinheim, 2005).
- Dowhan, W. Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu. Rev. Biochem.* **66**, 199–232 (1997).
- Zampighi, G., Simon, S. A., Robertson, J. D., McIntosh, T. J. & Costello, M. J. On the structural organization of isolated bovine lens fiber junctions. *J. Cell Biol.* **93**, 175–189 (1982).
- Kucerka, N. *et al.* Structure of fully hydrated fluid phase DMPC and DLPC lipid bilayers using X-ray scattering from oriented multilamellar arrays and from unilamellar vesicles. *Biophys. J.* **88**, 2626–2637 (2005).
- Palsdottir, H. & Hunte, C. Lipids in membrane protein structures. *Biochim. Biophys. Acta* **1666**, 2–18 (2004).
- Schey, K. L., Little, M., Fowler, J. G. & Crouch, R. K. Characterization of human lens major intrinsic protein structure. *Invest. Ophthalmol. Vis. Sci.* **41**, 175–182 (2000).
- Ball, L. E., Garland, D. L., Crouch, R. K. & Schey, K. L. Post-translational modifications of aquaporin O (AQP0) in the normal human lens: spatial and temporal occurrence. *Biochemistry* **43**, 9856–9865 (2004).
- Shiels, A. & Bassnett, S. Mutations in the founder of the MIP gene family underlie cataract development in the mouse. *Nature Genet.* **12**, 212–215 (1996).
- Shiels, A., Mackay, D., Bassnett, S., Al-Ghoul, K. & Kuszak, J. Disruption of lens fiber cell architecture in mice expressing a chimeric AQP0-LTR protein. *FASEB J.* **14**, 2207–2212 (2000).
- Francis, P. *et al.* Functional impairment of lens aquaporin in two families with dominantly inherited cataracts. *Hum. Mol. Genet.* **9**, 2329–2334 (2000).
- Francis, P., Berry, V., Bhattacharya, S. & Moore, A. Congenital progressive polymorphic cataract caused by a mutation in the major intrinsic protein of the lens, MIP (AQP0). *Br. J. Ophthalmol.* **84**, 1376–1379 (2000).
- Okamura, T. *et al.* Bilateral congenital cataracts result from a gain-of-function mutation in the gene for aquaporin-0 in mice. *Genomics* **81**, 361–368 (2003).
- Chepelinsky, A. B. The ocular lens fiber membrane specific protein MIP/aquaporin O. *J. Exp. Zool. A* **300**, 41–46 (2003).
- Lee, A. G. How lipids affect the activities of integral membrane proteins. *Biochim. Biophys. Acta* **1666**, 62–87 (2004).
- Jensen, M. O. & Mouritsen, O. G. Lipids do influence protein function—the hydrophobic matching hypothesis revisited. *Biochim. Biophys. Acta* **1666**, 205–226 (2004).
- Mitsuoka, K. *et al.* The structure of bacteriorhodopsin at 3.0 Å resolution based on electron crystallography: implication of the charge distribution. *J. Mol. Biol.* **286**, 861–882 (1999).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by NIH funding (to T.W.) and a Grant-in Aid for Specially Promoted Research (to Y.F.).

Author Information Coordinates and structure factors for junctional and non-junctional AQP0 have been deposited in the Protein Data Bank (accession codes 2B6O and 2B6P, respectively). Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.W. (twalz@hms.harvard.edu).

Creation of a six-atom 'Schrödinger cat' state

D. Leibfried¹, E. Knill¹, S. Seidelin¹, J. Britton¹, R. B. Blakestad¹, J. Chiaverini¹†, D. B. Hume¹, W. M. Itano¹, J. D. Jost¹, C. Langer¹, R. Ozeri¹, R. Reichle¹ & D. J. Wineland¹

Among the classes of highly entangled states of multiple quantum systems, the so-called 'Schrödinger cat' states are particularly useful. Cat states are equal superpositions of two maximally different quantum states. They are a fundamental resource in fault-tolerant quantum computing^{1–3} and quantum communication, where they can enable protocols such as open-destination teleportation⁴ and secret sharing⁵. They play a role in fundamental tests of quantum mechanics⁶ and enable improved signal-to-noise ratios in interferometry⁷. Cat states are very sensitive to decoherence, and as a result their preparation is challenging and can serve as a demonstration of good quantum control. Here we report the creation of cat states of up to six atomic qubits. Each qubit's state space is defined by two hyperfine ground states of a beryllium ion; the cat state corresponds to an entangled equal superposition of all the atoms in one hyperfine state and all atoms in the other hyperfine state. In our experiments, the cat states are prepared in a three-step process, irrespective of the number of entangled atoms. Together with entangled states of a different class created in Innsbruck⁸, this work represents the current state-of-the-art for large entangled states in any qubit system.

One promising candidate system for scalable universal quantum information processing (QIP) consists of atomic ions that are confined in electromagnetic traps and manipulated with laser beams⁹. Most of the basic ingredients for QIP¹⁰ have been demonstrated separately in the last few years in this system. Furthermore, some simple algorithms that could serve as primitives for larger scale QIP, including quantum error correction, teleportation, and the semiclassical quantum Fourier transform, have been implemented in the atomic-ion system.

Before large-scale QIP with atomic ions can become a reality, several challenges must be met successfully. In addition to building larger trap arrays and improving the classical control systems, it is necessary to demonstrate the ability to reliably create and maintain highly entangled states of many ions. Among such states, the so-called cat states, named after Schrödinger's cat¹¹, are of particular interest. Cat states are equal superpositions of two maximally different states (see below) and play a distinguished role in quantum information science. For three ion-qubits, they are also called Greenberger–Horne–Zeilinger (GHZ) states and provide a particularly clear demonstration of quantum non-locality⁶. In addition to the uses mentioned in the first paragraph, cat states can serve as a universal computation resource¹². They are also particularly sensitive benchmarks for demonstrating good control of quantum systems¹³ and the presence of entanglement. In the experiments described here, we prepared cat states of up to six ion-qubits with verifiable multi-particle entanglement.

The two states of a physical qubit are formally equivalent to the two states of a spin-1/2 magnetic moment in a magnetic field. Therefore we label our states $|\uparrow\rangle$ and $|\downarrow\rangle$ and define angular momentum operators $\hat{S}_x, \hat{S}_y, \hat{S}_z$ accordingly. In particular,

$\hat{S}_z|\uparrow\rangle = \frac{1}{2}|\uparrow\rangle$ and $\hat{S}_z|\downarrow\rangle = -\frac{1}{2}|\downarrow\rangle$ (for simplicity we set $\hbar = 1$). We define $|\uparrow, N\rangle \equiv |\uparrow\rangle_1|\uparrow\rangle_2\ldots|\uparrow\rangle_N$ and $|\downarrow, N\rangle \equiv |\downarrow\rangle_1|\downarrow\rangle_2\ldots|\downarrow\rangle_N$. In this notation, prototypical cat states of N qubits can be written as:

$$|N\text{ Cat}\rangle = \frac{1}{\sqrt{2}}(|\uparrow, N\rangle + e^{i\theta}|\downarrow, N\rangle) \quad (1)$$

To generate such states we initially prepare the ions in state $|\downarrow, N\rangle$ and then apply the following unitary operation to transform the initial state into $|N\text{ Cat}\rangle$ (ref. 7):

$$U_N = \left(\exp\left[i\frac{\pi}{2}J_x\right] \exp\left[i\frac{\xi\pi}{2}J_z\right] \right) \left(\exp\left[i\frac{\pi}{2}J_z\right] \right) \left(\exp\left[i\frac{\pi}{2}J_x\right] \right) \quad (2)$$

The operators in the left and right pairs of parentheses represent a common rotation by angle $\frac{\pi}{2}$ of all N qubits, written in terms of the global angular momentum operators composed of the sum of the N individual spin-1/2 operators $\vec{J} = \sum_{j=1}^N \vec{S}_j$ (Dicke operators). The operator in the middle pair of parentheses represents a global entangling interaction that is diagonal in the measurement basis spanned by all product states of N qubits, each in either $|\uparrow\rangle$ or $|\downarrow\rangle$, and can be implemented by generalizing the phase-gate mechanism described in ref. 14 (see also below). If N is odd, $\xi = 1$; $\xi = 0$ otherwise.

Because of experimental imperfections, we need a measure to indicate how close the generated state $|\Psi_N\rangle$ is to the ideal state $|N\text{ Cat}\rangle$. The simplest measure, called the fidelity, is the square modulus of the overlap of these two states:

$$F_{N\text{ Cat}} = |\langle\Psi_N|N\text{ Cat}\rangle|^2 \quad (3)$$

To determine the fidelity of $|\Psi_N\rangle$ it is sufficient to know the probabilities $P_{\uparrow N}, P_{\downarrow N}$ of being in $|\uparrow, N\rangle$ or $|\downarrow, N\rangle$ and the coefficient $C_{\downarrow N; \uparrow N}$ of the $|\downarrow, N\rangle \langle \uparrow, N|$ component of the density matrix¹⁵

$$F_{N\text{ Cat}} = \frac{1}{2}(P_{\uparrow N} + P_{\downarrow N}) + |C_{\downarrow N; \uparrow N}| \geq 2|C_{\downarrow N; \uparrow N}| \quad (4)$$

where the last inequality follows from the positive semidefiniteness of density matrices.

In general, the fidelity is not sufficient as a characterization of the entanglement properties of $|\Psi_N\rangle$. For $N > 2$ there is no single measure that quantifies entanglement in all circumstances because there are many different ways in which N qubits can be entangled¹⁶. Consider a partition of the N qubits into disjoint subsets A and B . The qubits of A are said to be unentangled with the qubits of B if the state is a product of a state of A and of another state of B , or if the state is a mixture of such product states. In this case, non-classical correlations between A and B are absent. Conversely, the N qubits are said to exhibit genuine N -particle entanglement if there is no partition into non-empty subsets A and B for which the state is unentangled. The entanglement of two N -qubit states $|\phi\rangle$ and $|\psi\rangle$ can be compared if it is possible to obtain $|\psi\rangle$ from $|\phi\rangle$ by distributing the qubits to different parties who then apply arbitrary quantum operations and communicate classically. This means of transforming one N -qubit

¹National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80305, USA. †Present address: Los Alamos National Laboratory P-21, MS D454, Los Alamos, New Mexico 87545, USA.

state to another is known as ‘local operations and classical communication’ (LOCC)¹⁷. For $N > 2$ there are N -particle entangled states that cannot be transformed into each other, even if $|\psi\rangle$ need only be obtained from $|\phi\rangle$ with non-zero probability of success¹⁸. This implies that such states belong to different entanglement classes and that it is not possible to compare the amount of entanglement by using LOCC transformations. One of the classes with genuine N -particle entanglement is characterized by the cat states.

Establishing genuine N -particle entanglement in the experiment requires making measurements that clearly distinguish the produced state from any incompletely entangled state. One approach is based on so-called entanglement witness operators¹⁹. If an entanglement witness operator has a negative expectation value for a state, then that state is definitely N -particle entangled. Thus, one way to experimentally determine the presence of entanglement is to measure the expectation value of a well-chosen witness and show that it is negative with sufficient statistical significance. N -particle entanglement of cat states can be proved with a particularly simple witness operator (derived from ref. 15) based on the projector onto the ideal state $|N \text{ Cat}\rangle$:

$$W = 1 - 2|N \text{ Cat}\rangle\langle N \text{ Cat}| \quad (5)$$

The expectation value of this operator is directly related to the fidelity in equation (3):

$$\langle W \rangle = 1 - 2F_{N \text{ Cat}} \leq 1 - 4|C_{|N;|N}| \quad (6)$$

A nice feature of this entanglement measure is that if it is negative, then copies of $|\Psi_N\rangle$ can be purified by LOCC to nearly pure cat states by means of a simple and robust purification procedure²⁰.

Another strategy for proving N -particle entanglement is based on a ‘depolarization’ method. Using LOCC operations, the original density matrix can be transformed into a standard, partially depolarized form in which N -particle entanglement becomes obvious²¹. Therefore, if the depolarized state is N -particle entangled then so is the original state. The depolarized state is definitely N -particle entangled if

$$2|C_{|N;|N}| > \max_j(P_j + P_{\bar{j}}) \quad (7)$$

where P_j is the probability that state $|j\rangle$ is found upon measurement, $|j\rangle$ denotes a sequence of N ions in state $|\uparrow\rangle$ or $|\downarrow\rangle$, not all states equal, and $|\bar{j}\rangle$ corresponds to $|j\rangle$ with $\uparrow$ and $\downarrow$ interchanged. The quantities $|C_{|N;|N}|$ and $P_j + P_{\bar{j}}$ are not changed by the depolarization method. Consequently, we can obtain them directly from observations of the prepared state without actually implementing the depolarization. Inequality (7) can be satisfied even if $F < 0.5$, but the purification process to obtain nearly pure cat states with LOCC from multiple copies may no longer be simple. For our experimental six-qubit cat state (below), since $\langle W \rangle$ was found to be only slightly negative, we used the depolarization method to conclusively establish that it was six-particle entangled.

By extending methods used previously to create two- and three-qubit entangled states^{7,14}, we entangled up to six ions in states that approximate N -qubit cat states. We confined $^9\text{Be}^+$ ions to the axis of a linear Paul trap with axial centre-of-mass (COM) frequencies between $\omega_{\text{COM}}/(2\pi) = 2.6$ MHz and $\omega_{\text{COM}}/(2\pi) = 3.4$ MHz and radial COM frequencies of approximately 8 MHz (ref. 22). All N axial motional modes of the ions were cooled to the ground state by extending the method of ref. 23. We prepared the internal state of each ion in the $|F = 2, m_F = -2\rangle \equiv |\downarrow\rangle$ hyperfine ground state by optical pumping, where F and m_F are the total angular momentum and the component of the angular momentum along the quantization axis. We take $|F = 1, m_F = -1\rangle \equiv |\uparrow\rangle$ as the other qubit state. The encoding operations U_N were realized using two-photon stimulated Raman transitions uniformly applied to all ions, incorporating a phase gate G_N that is an extension of the gate described in ref. 14. The phase gate was implemented by two laser beams that uniformly

illuminate all ions with a relative detuning $\omega_{\text{COM}} + \delta$ with $\delta \ll \omega_{\text{COM}}$, exerting a state-dependent axial optical dipole force on the ions¹⁴. The spacing of the ions was chosen such that the force was proportional to $\langle J_z \rangle$ (see Methods). If this dipole force is applied for a duration $t_g = 2\pi/\delta$, the motion of the COM mode is excited and de-excited in such a way that each state on which a non-zero net force acts ($\langle J_z \rangle \neq 0$) traverses a circle in phase space¹⁴ and acquires a phase given by the area circumscribed in phase space. This area is proportional to the square of the net force and therefore to $\langle J_z^2 \rangle$ for that state. In the experiment, the strength and detuning were adjusted to yield a phase of $\frac{\pi}{2} \langle J_z^2 \rangle$ on each component of the wavefunction, thus realizing the third (middle) operator in equation (2). For $N = 5$ (and therefore $\xi = 1$) the left operator (up to trivial phases) was realized by an appropriate change of the final pulse: $\exp[i\frac{\pi}{2}J_x] \exp[i\frac{\pi}{2}J_z] \rightarrow \exp[i\frac{\pi}{2}J_y]$.

After creating each cat state, we determined the populations in substates with equal numbers of $|\uparrow\rangle$ components by observing state-dependent fluorescence (see Fig. 1 and Methods). The most important information on the quality of the states resides in the magnitude of coherence, $C_{|N;|N|}$. A lower bound on this quantity can be extracted by executing the interferometry algorithm described in ref. 7 on the states. Here $|\Psi_N\rangle$ is ‘decoded’ by applying the operation $U_{N,\phi}$ to it. $U_{N,\phi}$ differs from U_N only by replacing the two J_x operations by operations $J_\phi = \cos(\phi)J_x + \sin(\phi)J_y$. Note that the effect of applying $U_{N,\phi}$ is equivalent to applying $\exp[-iJ_z\phi]U_N\exp[iJ_z\phi]$. Since this is followed by measurement, the final $\exp[-iJ_z\phi]$ has no observable effect. The first operator, $\exp[iJ_z\phi]$, is a rotation around the z -axis of the Bloch sphere and uniquely ‘labels’ the coherence between $|\downarrow, N\rangle$ and $|\uparrow, N\rangle$ with a phase that evolves at N times ϕ (ref. 24). The net effect is that $U_{N,\phi}$ transfers the coherence into a population difference. Ideally⁷

$$\begin{aligned} |\Psi_\phi\rangle &= U_{N,\phi}|N \text{ Cat}\rangle \\ &= -i \sin\left(\frac{N}{2}\phi\right) |\downarrow, N\rangle + i^{N+1+\xi} \cos\left(\frac{N}{2}\phi\right) |\uparrow, N\rangle \end{aligned} \quad (8)$$

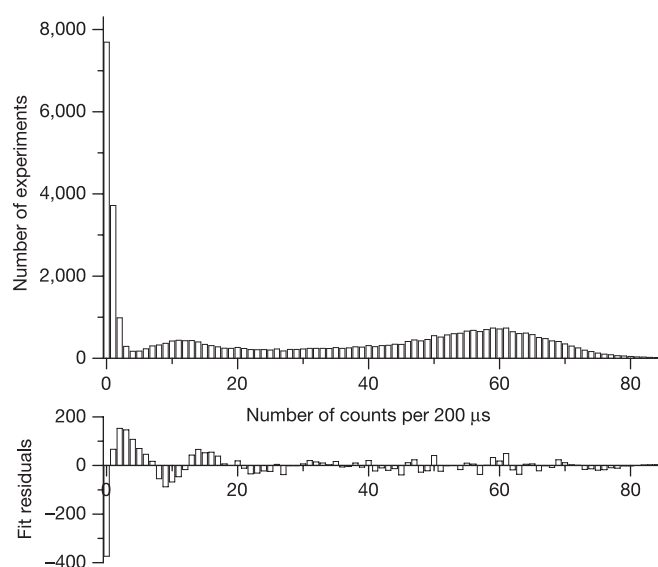


Figure 1 | Histogram and residuals of the $|6\text{Cat}\rangle$. Upper plot, the experimental histogram is fitted to a sum of seven Poissonian distributions with mean values corresponding to 0, 1, 2, ..., 6 ions in state $|\downarrow\rangle$. The fit yields the populations P_{16} , P_{16} , and upper bounds on all other populations (see Methods). The residuals of this fit are displayed in the lower plot. The sum of all positive residuals is 1,243.6 and is forced to be equal to the sum of all negative residuals by the fitting method. The deviations of the fit from the experimental distributions are mostly due to repumping of $|\uparrow\rangle$ into fluorescing states and non-uniformity of the fluorescence detection over the ion string.

In the experiment we observed a fluorescence signal, which has a contribution that oscillates N -times sinusoidally as ϕ is ramped through 2π . This oscillation can only arise from the coherence between $|\downarrow, N\rangle$ and $|\uparrow, N\rangle$ between applications of U_N and $U_{N,\phi}$. It has a maximal amplitude $A_{\max}$ determined by the difference in the fluorescence signal with all ions in state $|\downarrow\rangle$ and the signal with all ions in state $|\uparrow\rangle$, which is achieved if all the operations are implemented perfectly. In the imperfect case, the actual amplitude of the oscillation A yields a lower bound on $|C_{\downarrow N; \uparrow N}|$ via the inequality $|C_{\downarrow N; \uparrow N}| \geq \frac{1}{2}A/A_{\max}$. Because the imperfections in $U_{N,\phi}$ match those of U_N up to the high accuracy with which pulse phases are controlled, this bound is overly pessimistic. However, it cannot be improved without introducing additional assumptions.

Figure 2 shows the fluorescence as a function of phase ϕ for 4, 5 and 6 ions obtained by the method of the previous paragraph. As the number of ions increases, the coherence is more strongly affected by several sources of imperfection. Most importantly, decoherence due to spontaneous emission increases in proportion to the number of ions and the duration of the encoding operation U_N (for six ions we estimate an 18% probability of spontaneous emission per gate). In addition, the susceptibility to magnetic field noise, which washes out the fringe contrast, grows in proportion to N (for the noise observed in our laboratory environment we estimate a dephasing time of about $150\mu\text{s}$ for $|6\text{Cat}\rangle$ that has to be compared to about $50\mu\text{s}$ gate duration to implement U_6). As the length of the string of ions in the trap grows, rotations of all ions become less uniform owing to the approximately gaussian profile of the laser beams (for six ions, we estimate the spread in Rabi frequencies to be about 4%). Further

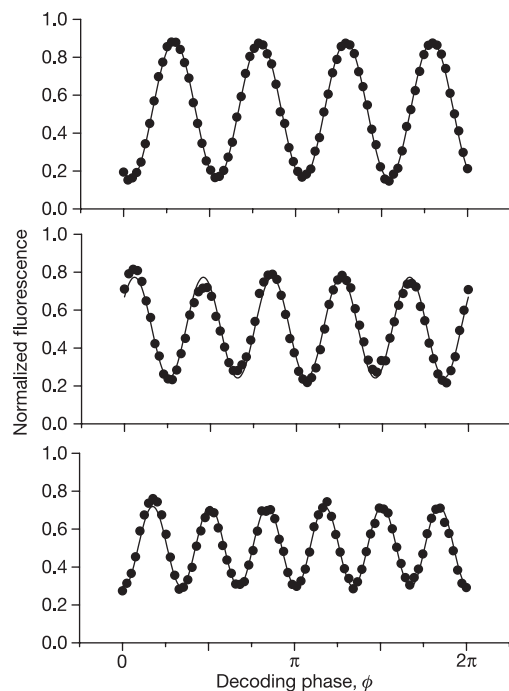


Figure 2 | Coherences of all prepared cat states. Measured traces (dots) of the normalized total ion fluorescence as the phase ϕ of the decoding gate $U_{N,\phi}$ is swept over a range of 2π for $N = 4, 5, 6$ (top to bottom). The fluorescence is normalized to the difference in the count rate of N ions all in state $|\downarrow\rangle$ and N ions all in $|\uparrow\rangle$. The contrast of the sinusoidal pattern with N oscillations is determined by a weighted least squares sin/cos decomposition restricted to frequencies of $0, 1, \dots, N$. The component with frequency N is shown (solid curves). The fitted contrast gives as lower bounds for the magnitudes of the coherences in the prepared states: $|C_{\downarrow 4; \uparrow 4}| \geq 0.698(3)/2$, $|C_{\downarrow 5; \uparrow 5}| \geq 0.527(3)/2$, $|C_{\downarrow 6; \uparrow 6}| \geq 0.419(4)/2$. The differences in the starting phase arise from additional J_z rotations inherent in our implementation of the phase gates. These rotations do not affect the character or fidelity of the produced states.

imperfections include laser beam intensity and phase fluctuations, both of the order of 5%. In spite of these imperfections, we achieved coherences of $|C_{\downarrow 4; \uparrow 4}| \geq 0.349(2)$, $|C_{\downarrow 5; \uparrow 5}| \geq 0.264(2)$, $|C_{\downarrow 6; \uparrow 6}| \geq 0.210(2)$. The error bars in these expressions are standard deviations obtained by resampling (see Methods). For four and five ions, the coherences alone are sufficient to prove genuine four- and five-particle entanglement (equation (6)). Together with the populations from poissonian fits (see Fig. 1 and Methods) we obtain fidelities $F_{4\text{Cat}} \geq 0.76(1)$, $F_{5\text{Cat}} \geq 0.60(2)$ and $F_{6\text{Cat}} \geq 0.509(4)$, all leading to negative values for the expectation value of witness operator W in equation (5): $\langle W_4 \rangle \leq -0.51(2)$, $\langle W_5 \rangle \leq -0.20(4)$ and $\langle W_6 \rangle \leq -0.018(8)$. Although $\langle W_6 \rangle$ is negative, it is not negative by a significant amount. To show that the state obtained has N -particle entanglement, we establish inequality (7). Although we cannot distinguish the populations P_j for different j with the same number of ions in state $|\uparrow\rangle$, we can place an upper bound on the quantities $(P_j + P_{\bar{j}})$ by using twice the maximum of the populations $P_{11}, \dots, P_{15}$, where P_{1k} is the probability that k ions are in the state $|\downarrow\rangle$. We obtained

$$|C_{\downarrow 6; \uparrow 6}| \geq 0.210(2) \geq \max(P_{1j}, j \in \{1, 2, 3, 4, 5\}) = 0.119(9) \quad (9)$$

so that the desired inequality is satisfied with high significance.

It should be possible to improve state preparation in future experiments. Decoherence due to spontaneous emission may be reduced by an appropriate choice of the Raman-detuning²⁵ or by an entirely different choice of ion species. Magnetic field noise may be suppressed by utilizing field-independent transitions²⁶. Fluctuations in beam power and phase may be actively cancelled by better feedback on the laser beams.

Cat states and the gates used to produce them are interesting in several respects. The state $|4\text{Cat}\rangle$ is one of the basic building blocks for the concatenated error correcting codes used in ref. 3 to achieve fault-tolerant quantum computing with realistically noisy devices. States with a higher number of qubits could be particularly valuable for more complicated fault-tolerant encoding schemes. The direct preparation method demonstrated here could significantly reduce the overhead in such schemes. The phase gate used in the production of $|4\text{Cat}\rangle$ has another interesting feature. Together with the phase gate described in ref. 14 and single-qubit rotations, it provides a universal gate set for quantum computing on a phase-decoherence-free subspace with logical qubits $\{|\downarrow\downarrow\rangle, |\uparrow\uparrow\rangle\}$ (refs 27, 28) where all gates are implemented with the same resources. Another notable feature of cat states is that they can be used to deterministically prepare Bell states of any two of the qubits by rotating and measuring the others and using the classical measurement outcomes to transform the two qubits. This feature is not shared by states such as W-states that typify some of the other entanglement classes. Bell states are the universal resource for quantum teleportation and communication between two parties. This feature is exploited in open-destination teleportation⁴. The multi-segmented trap architecture we are using should allow the distribution of entangled particles into separate locations to explore such protocols in future experiments.

Finally, spin-cat states are of particular interest in interferometry. If the contrast of the fringes in Fig. 2 were perfect, one could outperform the signal-to-noise limit of a perfect unentangled interferometer by a factor $\sqrt{N}$ and achieve the Heisenberg limit, the best possible signal-to-noise ratio within the limits of quantum uncertainty. Even with the imperfections of our experiments, all the states discussed in this Letter exhibit verified features that could not be reproduced with qubits that are not N -particle entangled, even with perfect experimental control.

METHODS

Inter-ion distance and relative phase of the dipole force. Phase gates for entangling N ions can be derived in a straightforward manner if the dipole force in the gate drive has the same phase on all ions⁷. For the arrangement of the beams in our experiment²³, the phase of the force repeats every 213 nm along the

alignment direction of the ions. The inter-ion distance is determined by the force equilibrium between mutual Coulomb repulsion of the ions and the confining force of the external trap potential, causing inter-ion distances to be unequal for $N > 3$. Nevertheless, a spacing where the dipole force has approximately the same phase for each ion can be achieved. As an example, the positions of four ions relative to the trap centre are $s(-1.437, -0.454, 0.454, 1.437)$, where $s = \sqrt[3]{e^2/(4\pi\epsilon_0 m_{\text{Be}} \omega_{\text{COM}}^2)}$ is a universal scaling parameter, with e the elementary charge, ϵ_0 the vacuum permittivity, m_{Be} the mass of the beryllium ion and ω_{COM} the axial COM frequency. We can take advantage of the fact that the two different distances have a ratio very close to an integer ratio, $1.437/0.454 - 19/6 \approx 0.0014$, and adjust the trap frequency ω_{COM} such that all four ions are spaced close to an integer number times 213 nm. The residual error would lead to a gate infidelity of 0.004, much smaller than the imperfection produced by other sources in our experiment. Similar considerations also hold for 5 and 6 ions. The problem could be completely overcome by driving the gate on a radial COM mode instead, an option that was not available for our current laser beam set-up.

Determination of populations from state dependent fluorescence. During one detection period (duration 200 μs) we typically detect on average $\lambda_0 \approx 0.5$ counts if all ions are projected into $|\uparrow\rangle$, and about $\lambda_1 \approx 10$ additional counts for each ion in state $|\downarrow\rangle$. The parameters λ_0 and λ_1 were derived by fitting mixtures of poissonian distributions to reference photon-count histograms obtained by running many experiments for each of a small number of states including $|\downarrow, N\rangle$. We used the maximum likelihood method for fitting the histograms and parametric-bootstrap resampling for determining standard errors in inferred quantities²⁹. Under the assumption that each ion fluoresces equally, the histograms should be well approximated by $E(P_{10} \text{ Poiss}(\lambda_0) + P_{11} \text{ Poiss}(\lambda_0 + \lambda_1) + \dots P_{1N} \text{ Poiss}(\lambda_0 + N\lambda_1))$, where E is the number of experiments contributing to the histogram and $\text{Poiss}(\lambda)$ is a poissonian distribution with mean λ . The parameters λ_0 and λ_1 determine the maximum possible amplitude of the phase oscillations in Fig. 2. This maximum amplitude is required for inferring a lower bound on $|C_{1N;1N}|$. The probabilities $P_{01}, \dots, P_{N1}$ were obtained by a maximum likelihood estimation of their values based on experimental population histograms obtained by direct observation of the prepared cat states. Up to 39,900 experiments were used to acquire these histograms. The desired witness expectations were computed according to the fits. From the populations P_{jk} , we determine the quantity $\frac{1}{2}(P_{1N} + P_{1N}) = \frac{1}{2}(P_{10} + P_{1N})$, the first term in equation (4). From the remaining populations, we can find upper bounds on P_j . For example ($N = 4$), $P_{1111} \leq P_{12}$. Figure 1 shows the measured histogram for the six-ion cat state together with the residuals between data and the fitted distribution. The histograms and fits for four and five ions look similar. From our fitted data we find (in the order $\{P_{10}, P_{11}, \dots, P_{1N}\}$) $P_{4\text{Cat}} = \{0.44(2), 0.079(2), 0.046(2), 0.063(2), 0.37(2)\}$, $P_{5\text{Cat}} = \{0.328(6), 0.143(6), 0.044(6), 0.033(6), 0.111(6), 0.340(6)\}$, $P_{6\text{Cat}} = \{0.317(9), 0.099(9), 0.061(9), 0.054(9), 0.068(9), 0.119(9), 0.282(9)\}$, for the relevant populations of the prepared cat states.

Received 27 July; accepted 20 September 2005.

- Shor, P. W. in *Proc. 37th Symp. on the Foundations of Computer Science (FOCS)* 56–65 (IEEE Press, Los Alamitos, California, 1996).
- Steane, A. M. & Iblinson, B. Fault-tolerant logical gate networks for css codes. Preprint at (<http://arxiv.org/quant-ph/0311014>) (2003).
- Knill, E. Quantum computing with realistically noisy devices. *Nature* **434**, 39–44 (2005).
- Zhao, Z. *et al.* Experimental demonstration of five-photon entanglement and open-destination teleportation. *Nature* **430**, 54–58 (2004).
- Hillery, M., Buzek, V. & Berthiaume, A. Quantum secret sharing. *Phys. Rev. A* **59**, 1829–1834 (1999).

- Greenberger, D. M., Horne, M., Shimony, A. & Zeilinger, A. Bell's theorem without inequalities. *Am. J. Phys.* **58**, 1131–1143 (1990).
- Leibfried, D. *et al.* Toward Heisenberg-limited spectroscopy with multiparticle entangled states. *Science* **304**, 1476–1478 (2004).
- Häffner, H. *et al.* Scalable multi-particle entanglement of trapped ions. *Nature* doi:10.1038/nature04279 (this issue).
- Cirac, J. & Zoller, P. Quantum computations with cold trapped ions. *Phys. Rev. Lett.* **74**, 4091–4094 (1995).
- DiVincenzo, D. The physical implementation of quantum computation. *Fortschr. Phys.* **48**, 771–783 (2000).
- Schrödinger, E. Die Gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* **23**, 807–812; 823–828; 844–849 (1935).
- Gottesman, D. & Chuang, I. L. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature* **402**, 390–393 (1999).
- Knill, E., Laflamme, R., Martinez, R. & Tseng, C.-H. An algorithmic benchmark for quantum information processing. *Nature* **404**, 368–370 (2000).
- Leibfried, D. *et al.* Experimental demonstration of a robust, high-fidelity geometric two ion-qubit phase gate. *Nature* **422**, 412–415 (2003).
- Sackett, C. A. *et al.* Experimental entanglement of four particles. *Nature* **404**, 256–259 (2000).
- Bennett, C. H., Popescu, S., Röhrlich, D., Smolin, J. A. & Thapliyal, A. V. Exact and asymptotic measures of multipartite pure state entanglement. *Phys. Rev. A* **63**, 012307 (2001).
- Bennett, C. H., Bernstein, J. J., Popescu, S. & Schumacher, B. Concentrating partial entanglement by local operations. *Phys. Rev. A* **53**, 2046–2052 (1996).
- Dür, W., Vidal, G. & Cirac, J. I. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A* **62**, 062314 (2000).
- Lewenstein, M., Kraus, B., Cirac, J. I. & Horodecki, P. Optimization of entanglement witnesses. *Phys. Rev. A* **62**, 052310 (2000).
- Dür, W., Aschauer, H. & Briegel, H.-J. Multiparticle entanglement purification for graph states. *Phys. Rev. Lett.* **91**, 107903 (2003).
- Dür, W. & Cirac, J. I. Multiparticle entanglement and its experimental detection. *J. Phys. A* **34**, 6837–6850 (2001).
- Rowe, M. A. *et al.* Transport of quantum states and separation of ions in a dual RF ion trap. *Quant. Inform. Comput.* **2**, 257–271 (2002).
- King, B. E. *et al.* Cooling the collective motion of trapped ions to initialize a quantum register. *Phys. Rev. Lett.* **81**, 1525–1528 (1998).
- Freeman, R. *Spin Choreography* (Oxford Univ. Press, Oxford, UK, 1998).
- Ozeri, R. *et al.* Hyperfine coherence in the presence of spontaneous photon scattering. *Phys. Rev. Lett.* **95**, 030403 (2005).
- Langer, C. *et al.* Long-lived qubit memory using atomic ions. *Phys. Rev. Lett.* **95**, 060502 (2005).
- Kielinski, D. *et al.* A decoherence-free quantum memory using trapped ions. *Science* **291**, 1013–1015 (2001).
- Bacon, D. M. *Decoherence, Control and Symmetry in Quantum Computers*. PhD thesis, Ch. 10 Univ. California, Berkeley (2001); (<http://arxiv.org/quant-ph/0305025>) (2003).
- Efron, B. & Tibshirani, R. J. *An Introduction to the Bootstrap* (Chapman & Hall, New York, 1993).

Acknowledgements This work was supported by the US National Security Agency (NSA) and Advanced Research and Development Activity (ARDA), by the Department of Defence Multidisciplinary University Research Initiative (MURI) programme administered by the Office of Naval Research and by NIST. We thank S. Glancy and J. Bollinger for comments on the manuscript. This paper is a contribution by the National Institute of Standards and Technology and not subject to US copyright.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.L. (dil@boulder.nist.gov).

Scalable multiparticle entanglement of trapped ions

H. Häffner^{1,3}, W. Hänsel¹, C. F. Roos^{1,3}, J. Benhelm^{1,3}, D. Chek-al-kar¹, M. Chwalla¹, T. Körber^{1,3}, U. D. Rapol^{1,3}, M. Riebe¹, P. O. Schmidt¹, C. Becher^{1,†}, O. Gühne³, W. Dür^{2,3} & R. Blatt^{1,3}

The generation, manipulation and fundamental understanding of entanglement lies at the very heart of quantum mechanics. Entangled particles are non-interacting but are described by a common wavefunction; consequently, individual particles are not independent of each other and their quantum properties are inextricably interwoven^{1–3}. The intriguing features of entanglement become particularly evident if the particles can be individually controlled and physically separated. However, both the experimental realization and characterization of entanglement become exceedingly difficult for systems with many particles. The main difficulty is to manipulate and detect the quantum state of individual particles as well as to control the interaction between them. So far, entanglement of four ions⁴ or five photons⁵ has been demonstrated experimentally. The creation of scalable multiparticle entanglement demands a non-exponential scaling of resources with particle number. Among the various kinds of entangled states, the ‘W state’^{6–8} plays an important role as its entanglement is maximally persistent and robust even under particle loss. Such states are central as a resource in quantum information processing⁹ and multiparty quantum communication. Here we report the scalable and deterministic generation of four-, five-, six-, seven- and eight-particle entangled states of the W type with trapped ions. We obtain the maximum possible information on these states by performing full characterization via state tomography¹⁰, using individual control and detection of the ions. A detailed analysis proves that the entanglement is genuine. The availability of such multiparticle entangled states, together with full information in the form of their density matrices, creates a test-bed for theoretical studies of multiparticle entanglement. Independently, ‘Greenberger–Horne–Zeilinger’ entangled states¹¹ with up to six ions have been created and analysed in Boulder¹².

We consider particles with the two levels $|S\rangle$ and $|D\rangle$. Then an N -particle W state

$$|W_N\rangle = (|D\cdots DDS\rangle + |D\cdots DSD\rangle + |D\cdots DSDD\rangle + \cdots + |SD\cdots D\rangle) / \sqrt{N} \quad (1)$$

consists of a superposition of N states where exactly one particle is in the $|S\rangle$ state while all others are in $|D\rangle$ (refs 6, 7). W states are N -particle entangled states of special interest: their entanglement is not only maximally persistent and robust under particle loss¹³, but also immune against global dephasing, and rather robust against bit flip noise. Because of this robustness, W states may lead to stronger non-classicality¹⁴ than GHZ states¹¹ for large numbers of particles. In addition, they may also be used for quantum communication^{15–17}.

The generation of such W states is performed in an ion-trap quantum processor¹⁸. We trap strings of up to eight $^{40}\text{Ca}^+$ ions in a linear Paul trap. Superpositions of the $S_{1/2}$ ground state and the metastable $D_{5/2}$ state of the Ca^+ ions (lifetime of the $|D\rangle$ level:

$\tau \approx 1.16$ s) represent the qubits. Each ion qubit in the linear string is individually addressed by a series of tightly focused laser pulses on the $|S\rangle \equiv S_{1/2}(m_j = -1/2) \leftrightarrow |D\rangle \equiv D_{5/2}(m_j = -1/2)$ quadrupole transition employing narrowband laser radiation near 729 nm. Doppler cooling on the fast $S \leftrightarrow P$ transition (lifetime ~ 8 ns) and subsequent sideband cooling prepare the ion string in the ground state of the centre-of-mass vibrational mode¹⁸. Optical pumping initializes the ions’ electronic qubit states in the $|S\rangle$ state. After preparing an entangled state with a series of laser pulses, the quantum state is read out with a CCD camera using state selective fluorescence¹⁸.

The W states are efficiently generated by sharing one motional quantum between the ions with partial swap operations (see Table 1)⁸. For an increasing number of ions, however, the initialization of the quantum register becomes more and more difficult as technical imperfections—like incomplete optical pumping—add up for each ion. Therefore, for $N = 6, 7, 8$, we first prepare the state $|0, DD\cdots D\rangle$ with $N\pi$ pulses on the carrier transition¹⁸, where the 0 refers to the motional state of the centre-of-mass mode. Then, laser light resonant with the $S \leftrightarrow P$ transition projects the ion string on the measurement basis. Absence of fluorescence indicates that all ions are prepared in $|D\rangle$. Similarly, we test the motional state with a single π pulse on the blue sideband¹⁸. Absence of fluorescence during a subsequent detection period indicates ground state occupation. Success of both checks (total success rate ≥ 0.7) confirms that the desired initial state $|0, DD\cdots D\rangle$ is indeed prepared. We can then start with the actual entangling procedure (step (1) in Table 1) and create $|W_N\rangle$ states ($N \leq 8$) in about 500–1,000 μs .

Full information of the N -ion entangled state is obtained via quantum state reconstruction by expanding the density matrix in a basis of observables¹⁹ and measuring the corresponding expectation values. In order to do this, we employ additional laser pulses to rotate the measurement basis prior to state detection¹⁰. We use 3^N different bases and repeat the experiment at least 100 times for each basis. For $N = 8$, this amounts to $\geq 656,100$ experiments and a total measurement time of 10 hours. To obtain a positive semi-definite density matrix ρ , we follow the iterative procedure outlined in ref. 20 for performing a maximum-likelihood estimation of ρ . The reconstructed density matrix for $N = 8$ is displayed in Fig. 1. To retrieve the fidelity $F = \langle W_N | \rho | W_N \rangle$, we adjust the local phases such that F is maximized (see Methods). The local character of those transformations implies that the amount of entanglement present in the system is not changed. We obtain fidelities $F_4 = 0.85$, $F_5 = 0.76$, $F_6 = 0.79$, $F_7 = 0.76$ and $F_8 = 0.72$ for the 4-, 5-, 6-, 7- and 8-ion W states, respectively.

The probabilistic nature of the measurement process requires an infinite number of measurements for a perfect reconstruction of the density matrix. In order to assess the error introduced by the finite number of measurements (quantum projection noise), we have used a Monte Carlo simulation to create up to 100 comparable data sets.

¹Institut für Experimentalphysik, ²Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 25, A-6020 Innsbruck, Austria. ³Institut für Quantenoptik und Quanteninformation der Österreichischen Akademie der Wissenschaften, Technikerstraße 21a, A-6020 Innsbruck, Austria. [†]Present address: Fachrichtung Technische Physik, Universität des Saarlandes, Postfach 151150, D-66041 Saarbrücken, Germany.

Table 1 | Creation of a $|W_N\rangle$ -state ($N = \{6,7,8\}$)

	Initialization		Entanglement
(i1)	$ 0, SSS \dots S\rangle$	(1)	$\xrightarrow{R_N^+(\arccos(1/\sqrt{N}))}$
	$\xrightarrow{R_N^+(\pi) R_{N-1}^+(\pi) \dots R_1^+(\pi)}$		$\frac{1}{\sqrt{N}} 0, SDD \dots D\rangle + \frac{\sqrt{N-1}}{\sqrt{N}} 1, DDD \dots D\rangle$
	$ 0, DDD \dots D\rangle$	(2)	$\xrightarrow{R_{N-1}^+(\arcsin(1/\sqrt{N-1}))}$
(i2)	Check state via fluorescence		$\frac{1}{\sqrt{N}} 0, SDD \dots D\rangle + \frac{1}{\sqrt{N}} 0, DSD \dots D\rangle + \frac{\sqrt{N-2}}{\sqrt{N}} 1, DDD \dots D\rangle$
	$\xrightarrow{R_1^+(\pi)}$	$\vdots$	$\vdots$
	$ 0, DDD \dots D\rangle$		$\frac{1}{\sqrt{N}} 0, SDD \dots D\rangle + \frac{1}{\sqrt{N}} 0, DSD \dots D\rangle + \dots + \frac{1}{\sqrt{N}} 1, DDD \dots D\rangle$
(i3)	Check state via fluorescence	(N)	$\xrightarrow{R_1^+(\arcsin(1/\sqrt{N}))}$
	$\xrightarrow{R_N^+(\pi)}$		$\frac{1}{\sqrt{N}} 0, SDD \dots D\rangle + \frac{1}{\sqrt{N}} 0, DSD \dots D\rangle + \dots + \frac{1}{\sqrt{N}} 0, DDD \dots S\rangle$
	$ 0, SDD \dots D\rangle$		

(i1)–(i3) are initialization steps; (1)–(N) are entanglement steps. First we initialize the ions via sideband cooling and optical pumping in the $|0, SSS \dots S\rangle$ state, where we use the notation $|n, x_N x_{N-1} \dots x_1\rangle$. n describes the vibrational quantum number of the ion motion and x_i their electronic state. We then prepare the $|0, DDD \dots D\rangle$ state with N π -pulses on the carrier transition applied to ions 1 to N , denoted by $R_i^+(\theta = \pi)$ (the notation is detailed in ref. 29; we do not specify the phase of the pulses because their particular value is irrelevant in this context). Then this state is checked for vanishing fluorescence with a photomultiplier tube. The same is done after trying to drive a π pulse on the blue sideband on ion 1 to ensure that the ion crystal is in the motional ground state. After this initialization, we transform the state to $|0, SDD \dots D\rangle$ with a carrier π pulse and start the entanglement procedure in step (1). This is carried out by moving most of the population to $|1, DDD \dots D\rangle$ with a blue sideband pulse of length $\theta_n = \arccos(1/\sqrt{n})$ leaving the desired part back in $|0, SDD \dots D\rangle$. Finally, we use $N - 1$ blue sideband pulses ($R_i^+(\theta_n)$) of pulse length $\theta_n = \arcsin(1/\sqrt{n})$ such that at each step we split off a certain fraction of the wave packet. Note that for an ion string in the ground state, blue-sideband pulses acting on an ion in the D state have no effect. For $N = \{4, 5\}$ we do not check the fluorescence, combine steps (i1) and (i3) and omit step (i2).

These data have been generated assuming ideal measurements on the reconstructed density matrix and using the measurement settings of the real experiment. For each of the artificial measurement sets a new density matrix was reconstructed via the maximum-likelihood method, and the spread of the expectation values of the observables was extracted.

For an investigation of the entanglement properties, we associate each particle k of a state ρ with a (possibly spatially separated) party A_k . We shall be interested in different aspects of entanglement between parties A_k , that is, the non-locality of the state ρ . A detailed entanglement analysis is achieved by investigating (1) the presence of genuine multipartite entanglement, (2) the distillability of multipartite entanglement and (3) entanglement in reduced states of two qubits.

First, we consider whether the production of a single copy of the state requires non-local interactions of all parties. This leads to the notion of multipartite entanglement and biseparability. A pure multipartite state $|\psi\rangle$ is called biseparable if two groups G_1 and G_2 within the parties A_k can be found such that $|\psi\rangle$ is a product state with respect to the partition

$$|\psi\rangle = |\chi\rangle_{G_1} \otimes |\eta\rangle_{G_2} \quad (2)$$

otherwise it is multipartite entangled. A mixed state ρ is called biseparable if it can be produced by mixing pure biseparable states $|\psi_i^{bs}\rangle$ —which may be biseparable with respect to different bipartitions—with some probabilities p_i , that is, the state can be written as $\rho = \sum_i p_i |\psi_i^{bs}\rangle \langle \psi_i^{bs}|$. If this is not the case, ρ is multipartite entangled. The generation of such a genuine multipartite entangled state requires interaction between all parties. In particular, a mixture of bipartite entangled states is not considered to be multipartite entangled. In order to show the presence of multipartite entanglement, we use the method of entanglement witnesses^{21–23}. An entanglement witness for multipartite entanglement is an observable with a positive expectation value on all biseparable states. Thus a negative expectation value proves the presence of multipartite entanglement. A typical witness for the states $|W_N\rangle$ would be²³:

$$\mathcal{W}_N = \frac{N-1}{N} \mathbb{1} - |W_N\rangle \langle W_N| \quad (3)$$

This witness detects a state as entangled if the fidelity of the W state exceeds $(N-1)/N$. However, more sophisticated witnesses can be constructed, if there is more information available on the state under

investigation than only the fidelity. To do so, we add other operators to the witness in equation (3) (see Methods) which take into account that certain biseparable states can be excluded on the grounds of the measured density matrix. Table 2 lists the expectation values for these advanced witnesses. The negative expectation values prove that in our experiment four-, five-, six-, seven- and eight-qubit entanglement has been produced.

Second, we consider the question of whether one can use many copies of the state ρ to distil one pure multipartite entangled state $|\psi\rangle$ by local means; that is, whether entanglement contained in ρ is qualitatively equivalent to multipartite pure state entanglement. For this aim one determines whether there exists a number M such that the transformation

$$\underbrace{\rho \otimes \rho \otimes \dots \otimes \rho}_{M \text{ copies}} \xrightarrow{\text{LOCC}} |\psi\rangle \quad (4)$$

is possible. Here, $|\psi\rangle$ is a multipartite entangled pure state (for

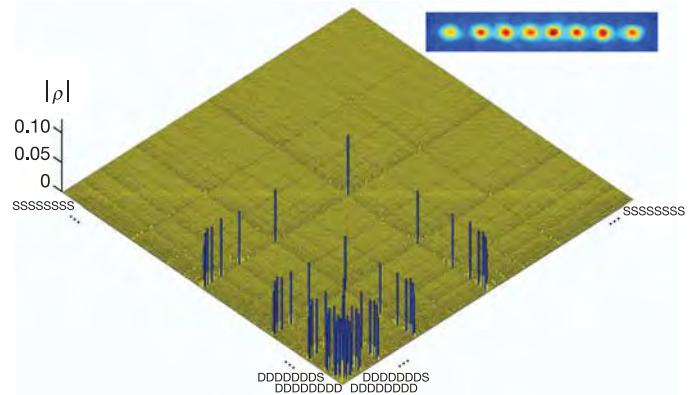


Figure 1 | Absolute values, $|\rho|$, of the reconstructed density matrix of a $|W_8\rangle$ state as obtained from quantum state tomography. DDDDDDDDD...SSSSSSSS label the entries of the density matrix ρ . Ideally, the blue coloured entries all have the same height of 0.125; the yellow coloured bars indicate noise. Numerical values of the density matrices for $4 \leq N \leq 8$ can be found in Supplementary Information. In the upper right corner a string of eight trapped ions is shown.

Table 2 | Entanglement properties of ρ_N

Property	$N = 3$	$N = 4$	$N = 5$	$N = 6$	$N = 7$	$N = 8$
F	0.824	0.846 (11)	0.759 (7)	0.788 (5)	0.763 (3)	0.722 (1)
$\text{tr}(\tilde{W}_N \rho_N)$	-0.532	-0.460 (31)	-0.202 (27)	-0.271 (31)	-0.071 (32)	-0.029 (8)
$\min(C_{kl})$	0.724	0.760 (34)	0.605 (23)	0.567 (16)	0.589 (9)	0.536 (8)
$\tilde{C}$	0.776	0.794 (23)	0.683 (15)	0.677 (11)	0.668 (5)	0.633 (3)
$\min(C'_{kl})$	0.294	0.229 (21)	0.067 (12)	0.049 (4)	0.035 (4)	0.022 (3)
$\tilde{C}'$	0.366	0.267 (12)	0.162 (6)	0.124 (3)	0.091 (2)	0.073 (1)

First row: fidelity after adjusting local phases (see Methods). Second row: expectation value of the witnesses $\tilde{W}_N$ (for $N = 8$, we additionally used local filters). Third and fourth row: respectively minimal and average concurrence between two qubits after observing the $|D\rangle$ state on the remaining $(N - 2)$ qubits. Fifth and sixth row: respectively minimal and average concurrence between two qubits after discarding the remaining $(N - 2)$ qubits. For completeness, we also analysed the data published previously in ref. 8 for $N = 3$.

example, $|\psi\rangle = |W_N\rangle$) and LOCC denotes a transformation using only local operations (with respect to the parties A_k) and classical communication. If such a transformation is possible, we call the state ρ multipartite distillable²⁴.

Technically, multipartite distillability follows from the possibility of generating maximally entangled singlet states $|\psi^-\rangle = (|DS\rangle - |SD\rangle)/\sqrt{2}$ between any pair of parties A_k, A_l by local means²⁴. The latter can readily be shown for all reconstructed density matrices. Performing measurements on all particles except k, l and restricting to outcomes $P_0 = |D\rangle\langle D|$ in all cases results in the creation of a two-qubit state ρ_{kl} . The density operator ρ_{kl} is distillable entangled if the concurrence C , a measure for two-qubit entanglement²⁵, is non-zero. This is the case for all k, l (see Table 2), which implies that ρ_N is multiparty distillable entangled. We remark that in practice one might use multiparticle entanglement purification protocols²⁶ to distil arbitrary entangled states.

Third, we investigate bipartite aspects of multiparticle entanglement²⁷, in particular the entanglement in the reduced states of two qubits. For W states this is of special interest, since for these states all reduced density operators of two particles are entangled, and the entanglement is in fact maximal^{6,28}. We investigate the bipartite entanglement by tracing out all but particles k, l and obtain the reduced density operators ρ'_{kl} . From these density matrices, we can now calculate the concurrence $C'_{kl} = C(\rho'_{kl})$ as a measure for the entanglement. For all N , we find that all reduced density operators are entangled (see Table 2). Note that the previous results (presence of multipartite entanglement and distillability) also imply that ρ is inseparable and in fact distillable with respect to any bipartition G_1-G_2 for all N .

Last, we address the scalability of our approach. Four major sources of deviations from the ideal W states are found: addressing errors, imperfect optical pumping, non-resonant excitations and frequency stability of the qubit-manipulation laser (see Methods). All of them are purely technical and thus do not represent fundamental obstacles to increasing the number of particles. We also note that the witnesses used to detect the multipartite entanglement do not require knowledge of the full density matrix. In particular, only $1 + 2N^2$ measurement settings are sufficient to determine the witnesses' expectation value²³. Thus the number of measurement settings does not increase exponentially with the number of particles. Also, the required blue sideband pulse area for a $|W\rangle$ state scales only with $\log N$ (see Table 1) while the time for a pulse with a given area is proportional to the square root of the ion crystal's mass, that is, to $\sqrt{N}$. Thus the overall favourable scaling behaviour of $\sqrt{N} \log N$ opens a way to study large-scale entanglement experimentally.

METHODS

Entanglement witness construction. Experimentally we do not create the W state given in equation (1), but rather a W state of the more general form

$$|\tilde{W}_N\rangle = (e^{i\varphi_1}|D\cdots DDS\rangle + e^{i\varphi_2}|D\cdots DSD\rangle + \dots + e^{i\varphi_{N-1}}|DSD\cdots D\rangle + \dots + e^{i(\varphi_N+\pi)}|SD\cdots DD\rangle)/\sqrt{N} \quad (5)$$

in which each ion has a different (local) phase φ_i . To determine the fidelity, we adjust these phases to maximize the overlap of the experimentally created W

state with $|\tilde{W}_N\rangle$. These small ($\varphi_i < 15^\circ$) phases appear because of a residual magnetic field gradient across the ion crystal and ac-Stark shifts induced by the laser pulses. Importantly, these effects are found to be constant and thus could be corrected for experimentally.

Witnesses for our experiment can be derived as follows: for N qubits we define the N states $|BS_i\rangle = |D\rangle_i \otimes |W_{N-1}\rangle$, which consist of $|D\rangle$ on the i th qubit and the state $|W_{N-1}\rangle$ on the remaining qubits. For the operator

$$\mathcal{Q} = \alpha |W_N\rangle\langle W_N| - \beta \sum_{i=1}^N |BS_i\rangle\langle BS_i| \quad (6)$$

we then compute the maximal expectation value for biseparable states. Since mixed biseparable states are convex combinations of pure biseparable states, it suffices to look at pure biseparable states, and thus we have to compute $\gamma = \max_{|\psi\rangle=|a\rangle\otimes|b\rangle} \langle\psi|\mathcal{Q}|\psi\rangle$ for all possible bipartitions²³. If we investigate a partition where $|a\rangle$ is a K -qubit state, it can be seen that the optimum $|a\rangle$ is of the form $|a\rangle = a_0|DD\cdots D\rangle + b_1|DD\cdots DS\rangle + b_2|D\cdots DSD\rangle + \dots + b_K|SDD\cdots D\rangle$. Then, from the matrix representation of $\mathcal{Q}$ one can deduce that the $a_0, b_1, \dots, b_K$ can be chosen real and finally that $b_i = b_j$ for all i, j . A similar statement can be proven for $|b\rangle$, thus for an arbitrary number of qubits the optimization procedure can be reduced to a four-parameter maximization with two normalization constraints, which can be efficiently solved numerically. The witness is then given by

$$\tilde{W}_N = \gamma 1_2 - \mathcal{Q} \quad (7)$$

where 1_2 denotes the identity operator on the space spanned by the elements of the computational basis which consists of $|S\rangle$ on at most two qubits. Adding the term $\gamma 1_2$ guarantees that $\tilde{W}_N$ is positive on all biseparable states. For the entanglement detection, we used the values $\alpha = 10$ and then $\beta = 2.98$, $\gamma = 2.2598$ for three qubits, $\beta = 2.87$, $\gamma = 0.8316$ for four qubits, $\beta = 2.35$, $\gamma = 0.3760$ for five qubits, $\beta = 1.94$, $\gamma = 0.1937$ for six qubits, $\beta = 1.638$, $\gamma = 0.1139$ for seven qubits, and $\beta = 1.4125$, $\gamma = 0.0764$ for eight qubits.

For $N = 8$ we have in addition optimized the witness using local filtering operations, that is, we applied a transformation $\tilde{W}_8^f = F\tilde{W}_8F^\dagger$ with $F = F_1 \otimes F_2 \otimes \dots \otimes F_8$. Here the F_i are operators acting on each qubit separately and are thus local operations. Therefore the new witness $\tilde{W}_8^f$ remains positive on all biseparable states. Finally, all witnesses have been normalized such that their expectation value for the maximally mixed state equals one and the local phases have been adjusted.

Experimental imperfections. For an investigation of the experimental imperfections, we simulate the preparation procedure by solving the Schrödinger equation with all relevant imperfections. This way, we identify four major sources of deviations from the ideal W states: addressing errors, imperfect optical pumping, non-resonant excitations, and laser frequency noise (including dephasing due to magnetic field noise). The trap frequency influences these experimental imperfections diametrically: for example, to keep the addressing error reasonably low (that is, less than 5%, where the addressing error is defined as the ratio of the Rabi frequencies between the addressed ion and the neighbouring ion(s)), we adjust the trap frequency such that the inter-ion distance in the centre of the ion string is about 5 μm . However, for large N the required reduction of the trap frequency implies that the sideband transition frequency moves closer to the carrier transition frequency. Thus the strong laser pulses driving the weak sideband transition cause more off-resonant excitations on the carrier transition, which in turn spoil the obtainable fidelity. Therefore we reduce the laser power for driving the sideband, which then results in longer preparation times and leads to an enhanced susceptibility to laser frequency noise. A compromise for the different ion numbers N is the following set of parameters: ($N = 4$: $\nu = 1.123$ MHz, $T_{2\pi} = 220$ μs), ($N = 5$: $\nu = 1.055$ MHz, $T_{2\pi} = 300$ μs), ($N = 6$: $\nu = 0.905$ MHz, $T_{2\pi} = 350$ μs), ($N = 7, 8$: $\nu = 0.813$ MHz, $T_{2\pi} = 380$ μs). Here ν is the trap frequency (centre of mass) and $T_{2\pi}$ is the time for a 2π pulse on the blue sideband. The fidelity reduction of $|W_6\rangle$ for the different imperfections are as follows: 0.1 (addressing

error), 0.07 (off-resonant excitations), 0.04 (laser frequency noise (200 Hz r.m.s.)). Another possible error source is imperfect ground state cooling. Intensity noise of the 729-nm laser ($\Delta I_{\max}/I \approx 0.03$) does not contribute significantly. Finally, we experimentally observed non-ideal optical pumping, which can result in a reduction of 0.02 of the fidelity per ion. In principle, this imperfection can be eliminated by single ion detection and subsequent π pulses on the carrier transition. Currently, for $N \geq 6$, we reduce the errors due to optical pumping and a part of the addressing errors by checking the initialization procedure with a detection sequence (see Table 1).

Received 19 July; accepted 3 October 2005.

- Schrödinger, E. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* **23**, 807–812 (1935).
- Schrödinger, E. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* **23**, 823–828 (1935).
- Schrödinger, E. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften* **23**, 844–849 (1935).
- Sackett, C. A. *et al.* Experimental entanglement of four particles. *Nature* **404**, 256–259 (2000).
- Zhao, Z. *et al.* Experimental demonstration of five-photon entanglement and open-destination teleportation. *Nature* **430**, 54–58 (2004).
- Dür, W., Vidal, G. & Cirac, J. I. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A* **62**, 062314 (2000).
- Zeilinger, A., Horne, M. A. & Greenberger, D. M. Higher-order quantum entanglement. *NASA Conf. Publ.* **3135**, 73–81 (1992).
- Roos, C. F. *et al.* Control and measurement of three-qubit entangled states. *Science* **304**, 1478–1480 (2004).
- Bennett, C. H. & DiVincenzo, D. P. Quantum information and computation. *Nature* **404**, 247–255 (2000).
- Roos, C. F. *et al.* Bell states of atoms with ultralong lifetimes and their tomographic state analysis. *Phys. Rev. Lett.* **92**, 220402 (2004).
- Greenberger, D. M., Horne, M. & Zeilinger, A. in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe* (ed. Kafatos, M.) 69–72 (Kluwer Academic, Dordrecht, 1989).
- Leibfried, D. *et al.* Creation of a six-atom 'Schrödinger cat' state. *Nature* doi:10.1038/nature 04251 (this issue).
- Briegel, H. J. & Raussendorf, R. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.* **86**, 000910 (2001).
- Sen(De), A. *et al.* Multiqubit W states lead to stronger nonclassicality than Greenberger–Horne–Zeilinger states. *Phys. Rev. A* **68**, 062306 (2003).
- Joo, J., Park, Y.-J., Lee, J., Jang, J. & Kim, I. Quantum secure communication via a W state. *J. Korean Phys. Soc.* **46**, 763–768 (2005).
- Joo, J., Lee, J., Jang, J. & Park, Y.-J. Quantum secure communication with W States. Preprint at (<http://arxiv.org/quant-ph/0204003>) (2002).
- Buhrman, H., van Dam, W., Høyer, P. & Tapp, A. Multiparty quantum communication complexity. *Phys. Rev. A* **60**, 2737–2741 (1999).
- Schmidt-Kaler, F. *et al.* How to realize a universal quantum gate with trapped ions. *Appl. Phys. B* **77**, 789–796 (2003).
- Fano, U. Description of states in quantum mechanics by density matrix and operator techniques. *Rev. Mod. Phys.* **29**, 74–93 (1957).
- Hradil, Z., Reháček, J., Fiurášek, J. & Ježek, M. Maximum-likelihood methods in quantum mechanics. *Lect. Notes Phys.* **649**, 59–112 (2004).
- Horodecki, M., Horodecki, P. & Horodecki, R. Separability of mixed states: Necessary and sufficient conditions. *Phys. Lett. A* **223**, 1–8 (1996).
- Terhal, B. M. Bell inequalities and the separability criterion. *Phys. Lett. A* **271**, 319–326 (2000).
- Bourennane, M. *et al.* Experimental detection of multipartite entanglement using witness operators. *Phys. Rev. Lett.* **92**, 087902 (2004).
- Dür, W. & Cirac, J. I. Classification of multiqubit mixed states: Separability and distillability properties. *Phys. Rev. A* **61**, 042314 (2000).
- Wootters, W. K. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.* **80**, 2245–2248 (1998).
- Miyake, A. & Briegel, H.-J. Distillation of multipartite entanglement by complementary stabilizer measurements. Preprint at (<http://arxiv.org/quant-ph/0506092>) (2005).
- Dür, W. & Cirac, J. I. Multiparticle entanglement and its experimental detection. *J. Phys. A* **34**, 6837–6850 (2001).
- Koashi, M., Bužek, V. & Imoto, N. Entangled webs: Tight bound for symmetric sharing of entanglement. *Phys. Rev. A* **62**, 050302 (2000).
- Gulde, S. *et al.* Implementing the Deutsch-Jozsa algorithm on an ion-trap quantum computer. *Nature* **421**, 48–50 (2003).

Supplementary information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge support by the Austrian Science Fund (FWF), by the European Commission (QGATES, CONQUEST, PROSECCO, QUPRODIS and OLAQUI networks), by the Institut für Quanteninformation GmbH, the DFG, and the ÖAW through project APART (W.D.). This material is based on work supported in part by the US Army Research Office. We thank P. Pham for the pulse modulation programmer, and A. Ostermann, M. Thalhammer and M. Ježek for help with the iterative reconstruction.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.H. (Hartmut.Haeffner@uibk.ac.at).

Origin of the metallic properties of heavily boron-doped superconducting diamond

T. Yokoya^{1,2}, T. Nakamura¹, T. Matsushita¹, T. Muro¹, Y. Takano³, M. Nagao³, T. Takenouchi⁴, H. Kawarada⁴ & T. Oguchi⁵

The physical properties of lightly doped semiconductors are well described by electronic band-structure calculations and impurity energy levels¹. Such properties form the basis of present-day semiconductor technology. If the doping concentration n exceeds a critical value n_c , the system passes through an insulator-to-metal transition and exhibits metallic behaviour; this is widely accepted to occur as a consequence of the impurity levels merging to form energy bands². However, the electronic structure of semiconductors doped beyond n_c have not been explored in detail. Therefore, the recent observation of superconductivity emerging near the insulator-to-metal transition³ in heavily boron-doped diamond^{4,5} has stimulated a discussion on the fundamental origin of the metallic states responsible for the superconductivity. Two approaches have been adopted for describing this metallic state: the introduction of charge carriers into either the impurity bands⁶ or the intrinsic diamond bands^{7–9}. Here we show experimentally that the doping-dependent occupied electronic structures are consistent with the diamond bands, indicating that holes in the diamond bands play an essential part in determining the metallic nature of the heavily boron-doped diamond superconductor. This supports the diamond band approach and related predictions, including the possibility of achieving dopant-induced superconductivity in silicon and germanium⁷. It should also provide a foundation for the possible development of diamond-based devices¹⁰.

An accurate description of the overall electronic structure is extremely important for understanding the mechanism of the superconductivity. We can approach this in two ways. First, we consider a localized picture that evolves into a metallic state in the vicinity of the insulator-to-metal transition. In diamond, carbon atoms are crystallized into a $2s2p^3$ three-dimensional network with covalent bonding. The resulting band structure¹¹ has a large valence bandwidth (22 eV); its top is located at the Γ point in the Brillouin zone (BZ) and is separated from the bottom of the conduction band with a 5.5-eV bandgap. For low carrier concentration, boron atoms probably replace the carbon sites by substitution and form an impurity level with an activation energy of 0.37 eV (ref. 12). As the boron concentration is increased, the wavefunctions of holes bound to an impurity site can overlap and the impurity level evolves into an impurity 'band'. The holes in a nearly localized 'band' are strongly affected by Coulomb forces, owing to small screening effects, and so electron–electron correlation is important. Both this and an extended s -wave superconductivity due to electron correlation were predicted in ref. 6. For the second approach, we consider the extended picture for the metallic states, starting from the diamond band structure. The doped holes lead to depopulation of the bands

with a dominant diamond character hybridized with boron states. Recent theoretical studies based on band structure calculations predicted that the superconductivity is driven by phonons strongly coupled to holes at the Γ point^{7–9}.

To distinguish the two scenarios we need to experimentally determine the band dispersion near the Fermi level E_F . To do this, we performed angle-resolved photoemission spectroscopy (ARPES) with the ability to measure band dispersions of crystalline solids. The intensity at E_F of normal-incidence ARPES spectra from a single-crystal diamond (111) film (BDD1) using photon energies from 770 to 870 eV were mapped with blue colour in Fig. 1a, which is a cross-section of the BZ for the diamond structure (Fig. 1b), including the Γ KLUX plane. The intensity distribution exhibits a maximum around the Γ point, consistent with the fact that diamond has a valence band maximum at the Γ point. Figure 1c shows an ARPES intensity map of BDD1 along the red curve in the Γ KLUX plane of the BZ (Fig. 1a) using a photon energy of 825 eV. ARPES using a soft X-ray (SXARPES) provides bulk sensitive band structures compared with vacuum ultraviolet ARPES¹³, but may have a larger window of momentum k parallel to the surface normal k_z , owing to larger available final states¹⁴.

In Fig. 1c, several dispersive features corresponding to experimentally obtained bands can be seen (denoted A–G). For the region of lower binding energy, we observe three dispersive features (A, B and C), all of which disperse towards E_F and appear to locate very close to E_F at Γ (we will discuss the band dispersion near E_F later in detail). Band C seems to disperse up to a binding energy of 14 eV and has an energy minimum at around $1.5 \text{ \AA}^{-1}$. At regions of higher binding energy, we observe higher-intensity features on both sides of Γ (D and E), suggesting a parabolic dispersion with a minimum at Γ . Band D seems to have a dispersion peak at around $1.5 \text{ \AA}^{-1}$, forming a bandgap with band C at a zone boundary. We also see an intensive feature (F) with nearly straight dispersion and a non-dispersive feature at 13 eV (G).

To compare the experimental bands with calculations, in Fig. 1d we show the same ARPES intensity map as in Fig. 1c superimposed with calculated band dispersions (solid white lines) along the red curve shown in Fig. 1a. The broken white lines are those calculated along the green line in Fig. 1a, shifted by the surface reciprocal lattice vector G . The calculated dispersions are energy-expanded by 10% and tentatively aligned so that the top of the valence band is located at 0.40 eV above E_F . We found that the experimental band dispersions (A, B, C, D and E) are very similar to the calculated band dispersions, including the bandgap at the zone boundary. Band F looks similar to the some of the broken lines and may thus be attributed to the surface umklapp band. The non-dispersive band G is not described with the

¹Japan Synchrotron Radiation Research Institute (JASRI)/SPRING-8, 1-1-1 Kouto, Sayo, Hyogo 679-5198, Japan. ²The Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Okayama 700-8530, Japan. ³National Institute for Materials Science, 1-2-1 Sengen, Tsukuba, Ibaraki 305-0047, Japan. ⁴School of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan. ⁵Department of Quantum Matter, Graduate School of Advanced Sciences of Matter (ADSM), Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima 739-8530 Japan.

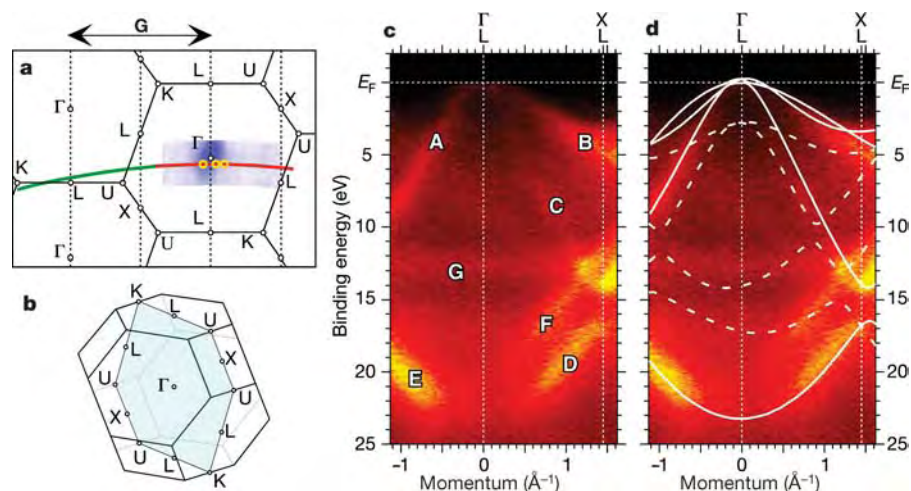


Figure 1 | Experimentally determined valence-band dispersions and the BZ of diamond. **a**, Two-dimensional BZ showing estimated measured k positions for the states near E_F (a red curve). The intensities near E_F from photon energy (770–870 eV) dependent SXARPES spectra from a single-crystal diamond (111) film are mapped with blue colour (higher spectral intensity corresponds to dense colour). The measured k positions are determined with a free-electron final-state model using:

$$(\hbar/2\pi)k_{\perp} = [2m(E\cos^2\theta + V_0)]^{1/2} - k_{\perp\text{photon}},$$

where $\hbar$ is Planck's constant, m is the electron mass, E and θ are the kinetic

energy and the polar angle of an emitted photoelectron, and $k_{\perp\text{photon}}$ is the surface normal momentum component of the photon. G is the surface reciprocal lattice vector. The yellow circles are k_F values determined from the MDC analysis. **b**, Three-dimensional BZ for the f.c.c. diamond crystal. The pale-blue plane includes the measured k positions shown in **a**. **c**, A valence-band SXARPES intensity map from a single-crystal diamond (111) film using a photon energy of 825 eV. **d**, Comparison of the same SXARPES intensity map with calculated band dispersions of diamond (white full and broken lines) along the red and green curves in **a**.

calculated dispersions, but may be related to the sharp s - p -band-derived density of states (DOS) observed from previous X-ray photoemission spectroscopy¹⁵, such as for the SXARPES of graphite¹⁶. These indicate that the gross electronic structure of diamond is retained with heavy boron doping.

As mentioned above, we find that the highly dispersive bands approach very close to E_F in Fig. 1c. To look more carefully at the states near E_F , we performed another SXARPES measurement for BDD1 with a smaller step size and with a higher signal-to-noise ratio, as shown in Fig. 2a. We now see the three bands near E_F more clearly and find that they disperse across E_F as is evident from a reduction of intensity along $k = 0$ near E_F . The band crossing can further be confirmed from the momentum distribution curve (MDC) at E_F and at binding energies of 0.3, 0.6 and 0.9 eV with an energy window of ± 0.15 eV. These MDCs have a multiple-peak structure that can be well reproduced with three lorentzian functions, as shown in Fig. 2b. The k positions of the centres of the lorentzian functions for MDC at E_F correspond to the Fermi momenta ($k_F^A = -0.15 \pm 0.04 \text{ \AA}^{-1}$, $k_F^C = 0.11 \pm 0.02 \text{ \AA}^{-1}$ and $k_F^B = 0.27 \pm 0.02 \text{ \AA}^{-1}$). These SXARPES results indicate that the observed diamond bands cross E_F forming hole pockets at the Γ point. The observation that the experimentally obtained band dispersions are similar to those calculated motivates us to determine the position of E_F with respect to the top of the valence band of the diamond band. We found that a location of E_F at 0.40 ± 0.2 eV below the top of the valence band, as shown in Fig. 2c, was necessary to relate k_F^A , k_F^B and k_F^C to our band structure calculations (white lines). This corresponds to the region of the carrier concentration: $6.6 \times 10^{20} \text{ cm}^{-3} < n = 1.9 \times 10^{21} \text{ cm}^{-3} < 4.4 \times 10^{21} \text{ cm}^{-3}$. These values are found to have a slightly lower value compared to the boron concentration from secondary ion mass spectroscopy (SIMS) measurements ($n_{\text{B,SIMS}} = 8.37 \times 10^{21} \text{ cm}^{-3}$).

To study the evolution of the states near E_F , we performed additional near- E_F SXARPES for BDD2 and BDD3 with lower boron concentrations, as shown in Fig. 2d and e. In Fig. 2d, we observe clearer dispersions in BDD2 than in BDD1, indicative of a sharpening of electronic structures in BDD2. But the band dispersions are similar to each other overall. From the MDC analysis for BDD2 (Fig. 2f), we obtained values for the Fermi

momenta ($k_F^A = -0.09 \pm 0.01 \text{ \AA}^{-1}$, $k_F^C = 0.06 \pm 0.008 \text{ \AA}^{-1}$ and $k_F^B = 0.19 \pm 0.007 \text{ \AA}^{-1}$). The smaller k_F values of BDD2 compared to those of BDD1 indicate that the number of carriers decreases with lowered boron concentration. We also found that a shift of 0.2 ± 0.1 eV is needed to match the band calculations to experimental k_F values, as shown in Fig. 2h. In Fig. 2e, the band dispersions of BDD3, which has a much lower boron concentration, are similar to those of BDD1 and BDD2 except near E_F . The results show a qualitative difference at the top of the band dispersion, where the plot for BDD3 has a local maximum at the Γ -point, while the plots for BDD1 and BDD2 show local minima at the Γ -point. This is consistent with the band structure calculation for E_F positioned just at the valence band maximum, as shown in Fig. 2i. This indicates that the bands of BDD3 do not cross E_F . We could not measure the band dispersion of pure diamond because of charging-up effects. However, the bandwidths estimated as a function of doping are $23.5 \text{ eV} \pm 0.5 \text{ eV}$ for both BDD1 and BDD3, indicating that a bandwidth change cannot be discerned, which is in agreement with the experimental bandwidth of undoped diamond ($23.0 \pm 0.2 \text{ eV}$) (ref. 17).

It may be interesting to estimate fundamental physical parameters from the MDC analysis and compare those values from samples with different boron concentrations. For band A of BDD1 and BDD2, the mean free path values λ , estimated from the half-width at half-maximum of the lorentzian for MDC at E_F are 5.3 and 8.6 $\text{\AA}$. From fits to the peak positions for the four binding energies (0.0 ($=E_F$), 0.3, 0.6 and 0.9 eV), we obtain Fermi velocities $|v_F|$ of 7.9 and 7.0 $\text{eV \AA}$, which correspond to values for the lifetime of the carrier τ ($=\lambda/v_F$) of 2.8 and 5.1 fs for BDD1 and BDD2. We note that because diamond is a three-dimensional material, contributions from the final-state electron lifetime cannot be neglected. Therefore the obtained λ and τ values give its lowest limit. Nonetheless, this simple analysis indicates a shorter lifetime of carriers in BDD1 than in BDD2.

These results indicate that the doped holes in this hole-concentration region enter into the top of the diamond valence band, accompanied by a shift of E_F . This justifies the band approach to the metallic states of heavily boron-doped diamond. This is

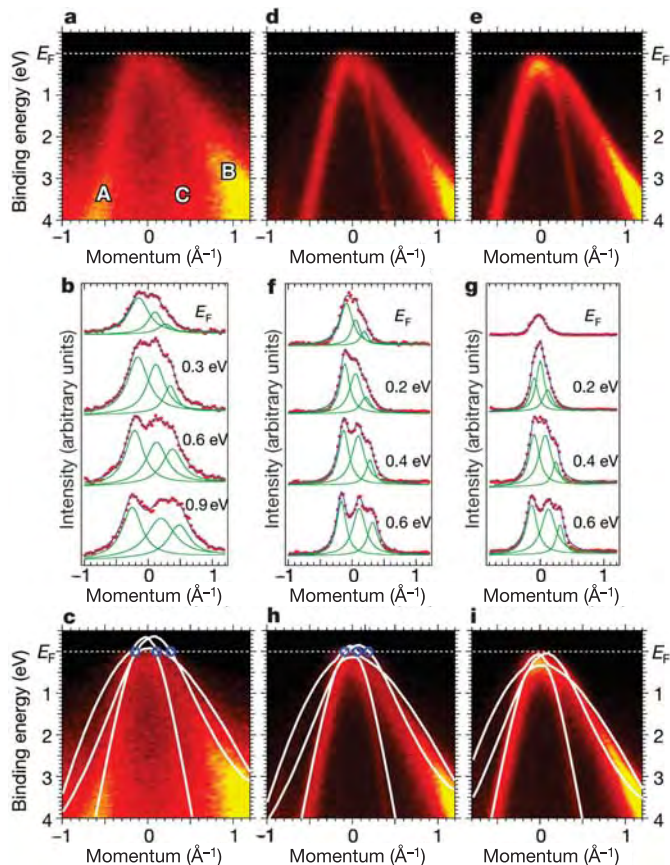


Figure 2 | Experimental band dispersions near E_F as a function of boron concentration and the results of MDC analysis. **a**, A near- E_F SXARPES intensity map from a single-crystal diamond (111) film (BDD1) using a photon energy of 825 eV along the red curve in Fig. 1a. **b**, MDCs at binding energies of 0.0 ($= E_F$), 0.3, 0.6 and 0.9 eV with an energy window of ± 0.15 eV (red dots) for BDD1. The blue line for each MDC is the sum of three lorentzians (green lines) corresponding to the three bands (A, B and C). **c**, A comparison of the experimental data with band structure calculations (white curves). k_F values determined from the MDC analysis are shown as blue circles. The white curves are the calculated band dispersion energies expanded 10% and shifted by 0.4 eV. That the calculated band dispersion splits at the top of the valence band corresponds to the measured k position, which is slightly away from the Γ point, as shown in Fig. 1a. The bands are indeed degenerate with the same valence-band maximum at the Γ point as for the calculated results of the Γ point. **d**, **e**, As for **a** but for BDD2 and BDD3, respectively. **f**, **g**, MDCs at binding energies of 0.0 ($= E_F$), 0.2, 0.4 and 0.6 eV with an energy window of ± 0.10 eV for BDD2 and BDD3, respectively, which emphasize the difference in MDC lineshapes in the near- E_F region. Descriptions for lines and symbols are the same as in **b**. Please note the very small intensity at E_F for BDD3 compared with the intensity at 0.2 eV for BDD3 as well as the intensity at E_F for BDD1 and BDD2, suggesting a qualitative difference in the E_F position. **h**, **i**, As for **c** but for BDD2 and BDD3, respectively. The energy-expanded calculated band dispersions are shifted by 0.2 eV for BDD2, but they are not shifted for BDD3.

different from the impurity band model⁶ and recent X-ray absorption spectroscopy (XAS) studies interpreting holes in the valence band located at about 1.3 eV below the valence band maximum, regardless of the doping level¹⁸. However, XAS measures site-projected, angular-momentum-projected unoccupied DOS in the presence of an attractive core-hole potential which can modify electronic energy levels compared to the single-particle DOS. In contrast, ARPES directly provides the energy- and momentum-resolved single-particle DOS with respect to E_F . The remarkable consistency with band dispersions observed in the present study as a function of doping confirms the validity of ARPES. We observed no additional

structure at a binding energy of 1.3 eV in the occupied DOS. In optimally doped high-temperature superconductors, the smaller bandwidth of the band crossing E_F compared to the band calculations has been reported and attributed to electron correlation effects¹⁹. This is in sharp contrast to heavily boron-doped diamond, where the experimental bandwidth is even slightly wider than that from calculations. This suggests a negligible role for the correlation effect in heavily boron-doped diamond.

The results revealed an overall occupied electronic structure directly related to the metallicity of the heavily boron-doped diamond superconductor. This provides experimental support for the very recent theoretical studies^{7–9} that adopt the band structure approach to understanding the electronic properties of doped diamond and related materials. An understanding of the electronic structure of heavily boron-doped diamond may also be important for developing diamond-based devices¹⁰ that make use of the unique properties of diamond.

METHODS

Homoepitaxially grown heavily boron-doped (111) diamond films (BDD1, BDD2 and BDD3) were made using a microwave plasma assisted chemical vapour deposition (CVD) method as described elsewhere^{5,20}. T_c of 7.0 K of BDD1 was confirmed by the measurement of the onset of magnetization after the present photoemission measurements. The T_c determined from the onset of magnetization is lower than that measured for the onset of resistivity and normally corresponds to zero resistivity. BDD2 and BDD3 did not show onset of magnetization above 1.7 K. However, BDD2 showed onset of resistivity measured below 2.5 K. SIMS measurements for films (BDD1, BDD2, BDD3) made under the same conditions as the samples we used gave boron concentrations of 8.37×10^{21} , 1.18×10^{21} and $2.88 \times 10^{20} \text{ cm}^{-3}$, respectively.

SXARPES measurements were performed at BL25SU, Spring-8, on a spectrometer built using a Scienta SES200 electron analyser. The energy and angular resolution were set to ~ 250 meV and $\pm 0.1^\circ$ (corresponding to $\pm 0.026 \text{ \AA}^{-1}$) for a photon energy of 825 eV, respectively, to obtain a reasonable count rate. Samples were cooled using a closed-cycle He refrigerator. Sample T was measured using a chromel-AuFe thermocouple mounted close to the sample. The base pressure of the spectrometer was better than 3×10^{-8} Pa. The sample orientation was measured *ex situ* using Laue photography. Location of the Γ point with respect to the measured ARPES data as well as the E_F position in the calculated spectra are determined by comparison between the experimental and calculated band dispersions near E_F . All measurements have been done for surfaces prepared with annealing at 400 °C to reduce oxygen contamination at the surface. E_F of the samples was referenced to that of a gold film evaporated onto the sample substrate measured just after the sample measurements. The measured k positions are determined with a free-electron final-state model by taking the photon momentum into consideration.

Band structure calculations for diamond are carried out within the local density approximation to the density functional theory. Kohn–Sham equations are solved self-consistently with the all-electron full-potential linear-augmented plane-wave (LAPW) method. The lattice constant of the face-centred cubic (f.c.c.) diamond structure is assumed to be 3.56 Å. The ‘muffin-tin’ sphere radius is taken as $R = 0.75 \text{ \AA}$. The LAPW basis functions up to $RK_{\text{max}} = 7.8$ are used to expand the wavefunctions.

Received 14 July; accepted 29 September 2005.

1. Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* Ch. 28, 561–587 (Saunders College, Fortworth, 1976).
2. Mott, N. *Metal-Insulator Transitions* Ch. 5, 145–169 (Taylor & Francis, London, 1990).
3. Bustarret, E., Gheeraert, F. & Watanabe, K. Dependence of the superconducting transition temperature on the doping level in single-crystalline diamond films. *Phys. Rev. Lett.* **93**, 237005 (2004).
4. Ekimov, E. A. *et al.* Superconductivity in diamond. *Nature* **428**, 542–545 (2004).
5. Takano, Y. *et al.* Superconductivity in diamond thin films well above liquid helium temperature. *Appl. Phys. Lett.* **85**, 2851–2853 (2004).
6. Baskaran, G. Resonating valence bond mechanism of impurity band superconductivity in diamond. Preprint at (<http://arXiv.org/cond-mat/0404286>) (2004).
7. Boeri, L., Kortus, J. & Andersen, O. K. Three-dimensional MgB_2 -type superconductivity in hole-doped diamond. *Phys. Rev. Lett.* **93**, 237002 (2004).
8. Lee, K.-W. & Pickett, W. E. Superconductivity in boron-doped diamond. *Phys. Rev. Lett.* **93**, 237003 (2004).

9. Blasé, X., Adessi, Ch. & Connetable, D. Role of the dopant in the superconductivity of diamond. *Phys. Rev. Lett.* **93**, 237004 (2004).
10. Nebel, C. E. & Ristein, J. *Thin-film Diamond II* Ch. 4–8, 121–358 (Elsevier, Amsterdam, 2004).
11. Zunger, A. & Freeman, A. J. Ground-state electronic properties of diamond in the local-density formalism. *Phys. Rev. B* **15**, 5049–5065 (1977).
12. Collins, A. T. & Williams, A. W. S. The nature of the acceptor centre in semiconducting diamond. *J. Phys. C* **4**, 1789–1800 (1971).
13. Kamakura, N. *et al.* Layer dependent band dispersion and correlation using soft X-ray ARPES. *Europhys. Lett.* **67**, 240–246 (2004).
14. Inglesfield, J. E. & Plummer, E. W. in *Angle-resolved Photoemission* (ed. Kevan, S. D.) Ch. 2, 15–61 (Elsevier, Amsterdam, 1992).
15. Gravell, R. G. *et al.* X-ray photoemission cross-section modulation in Diamond, Silicon, Germanium, Methane, Silane, and Germane. *Phys. Rev. B* **7**, 5313–5316 (1973).
16. Matsushita, T. *et al.* Angle-resolved soft X-ray photoemission for the valence band of graphite. *Surf. Rev. Lett.* **9**, 1321–1326 (2002).
17. Jimenez, I. *et al.* Accurate valence band width of diamond. *Phys. Rev. B* **56**, 7215–7221 (1997).
18. Nakamura, J. *et al.* Holes in the valence band of superconducting boron-doped diamond film studied by soft X-ray absorption and emission spectroscopy. Preprint at (<http://arXiv.org/cond-mat/0410144>) (2004).
19. Campuzano, J. C., Norman, M. R. & Randeria, M. *The Physics of Superconductors* Vol. II, Ch. 5, 184 (Springer, Berlin/Heidelberg, 2004).
20. Umezawa, H. *et al.* Advantage on superconductivity of heavily boron-doped (111) diamond films. Preprint at (<http://arXiv.org/cond-mat/0503303>) (2004).

Acknowledgements We thank A. Chainani for valuable discussions and critical reading of the manuscript. We thank N. Yamada and J. Nakamura for discussions. We thank T. Kinoshita for supporting our experimental plan to do a doping dependence study. We thank I. Sakaguchi for SIMS measurements. This study was supported by Grants-in-Aid for Young Scientists and for Exploratory Research from Japan Society for the Promotion of Science.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to T.Y. (yokoya@cc.okayama-u.ac.jp).

Chemically tailorable colloidal particles from infinite coordination polymers

Moonhyun Oh¹ & Chad A. Mirkin¹

Micrometre- and nanometre-sized particles play important roles in many applications, including catalysis¹, optics^{2,3}, biosensing^{4–8} and data storage⁹. Organic particles¹⁰ are usually prepared through polymerization of suitable monomers¹¹ or precipitation methods¹². In the case of inorganic materials, particle fabrication tends to involve the reduction of a metal salt¹³, or the controlled mixing of salt solutions supplying a metal cation and an elemental anion (for example, S^{2-} , Se^{2-} , O^{2-})¹⁴, respectively; in some instances, these methods even afford direct control over the shape of the particles produced^{14–17}. Another class of materials are metal-organic coordination polymers^{18–23}, which are based on metal ions coordinated by polydentate organic ligands and explored for potential use in catalysis¹⁸, gas storage^{19,20}, nonlinear optics²¹ and molecular recognition and separations^{22,23}. In a subset of these materials, the use of organometallic complexes as ligands (so-called metalloligands) provides an additional level of tailorability, but these materials have so far not yet been fashioned into nano- or microparticles. Here we show that simple addition of an initiation solvent to a precursor solution of metal ions and metalloligands results in the spontaneous and fully reversible formation of a new class of metal-metalloligand particles. We observe initial formation of particles with diameters of a few hundred nanometres, which then coalesce and anneal into uniform and smooth microparticles. The ease with which these particles can be fabricated, and the ability to tailor their chemical and physical properties through the choice of metal and organic ligand used, should facilitate investigations of their scope for practical applications.

We have discovered that spherical micro- and nanoparticles composed of polymerized metal-ligand networks can be made by the coordination-chemistry-induced assembly of metal ions and

homochiral carboxylate-functionalized binaphthyl bis-metallotridentate Schiff base (BMSB) building blocks (Fig. 1). In a typical experiment, a precursor pyridine solution consisting of a 1:1 mixture of the appropriate metal acetate salt ($M'(OAc)_2$, $M' = Zn, Cu$ and Ni , **1**) and BMSB, **2**, is prepared (alternatively, a 3:1 mixture of **1** and enantiopure carboxylate-functionalized binaphthyl bis-tridentate Schiff base (BSB) can be used). Slow addition of an initiation solvent such as diethyl ether or pentane results in the spontaneous formation of spherical inorganic polymer particles, **3**. These particles form via coordination of the carboxylate groups on the BMSB precursor with the metal ions supplied by the acetate salt, and the polymerization process is completely reversible, as evidenced by the formation of the starting materials upon addition of excess pyridine (Fig. 1).

The homochiral BMSB building blocks **2a–c** are key components which are readily polymerizable through their carboxylate groups; the choice of BMSB ligand, type of metallation, and ancillary ligands makes it possible to manipulate the chemical and physical properties of the resulting structures **3** in a systematic manner. The BMSB building blocks were prepared from the reaction of two equivalents of $M(OAc)_2$ ($M = Zn, Cu$ and Ni) and one equivalent of BSB in a *N,N*-dimethylformamide (DMF) solution at room temperature. The building blocks **2a–c** have been characterized by ¹H NMR, ¹³C NMR, infrared spectroscopy and electrospray ionization mass spectrometry, and all data are consistent with the proposed structures. Compound **2b** also was characterized in the solid state by a single-crystal X-ray diffraction study (Supplementary Information).

Coordination polymers formed from metal ions and carboxylate-functionalized building blocks are well known in transition metal coordination chemistry^{18–22}. They are typically prepared as macroscopic crystalline materials by one of several methods, including slow diffusion of solvent into solutions consisting of precursors,

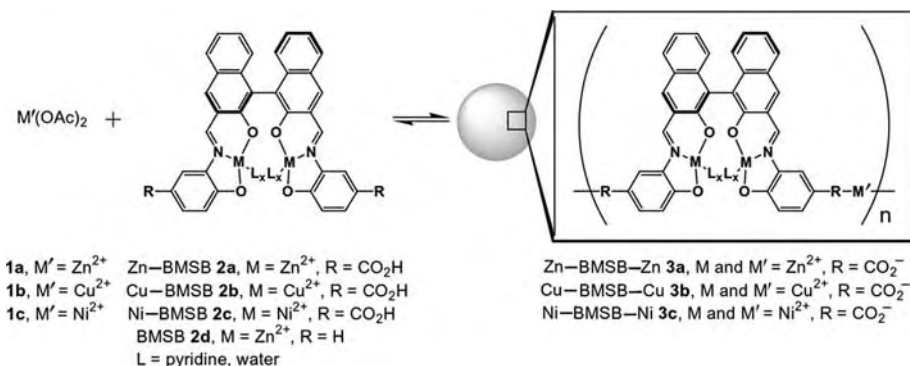


Figure 1 | Preparation of particles **3a–c**. Addition of diethyl ether or pentane into the reaction mixture containing **1** and **2** in pyridine results in particle formation and precipitation. Addition of excess pyridine dissolves

the particles and redisperses the ligand and metal ion building blocks. The reaction is not truly reversible as written because the M' ions can coordinate pyridine. BMSB, binaphthyl bis-metallotridentate Schiff base.

¹Department of Chemistry and Institute for Nanotechnology, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208-3113, USA.

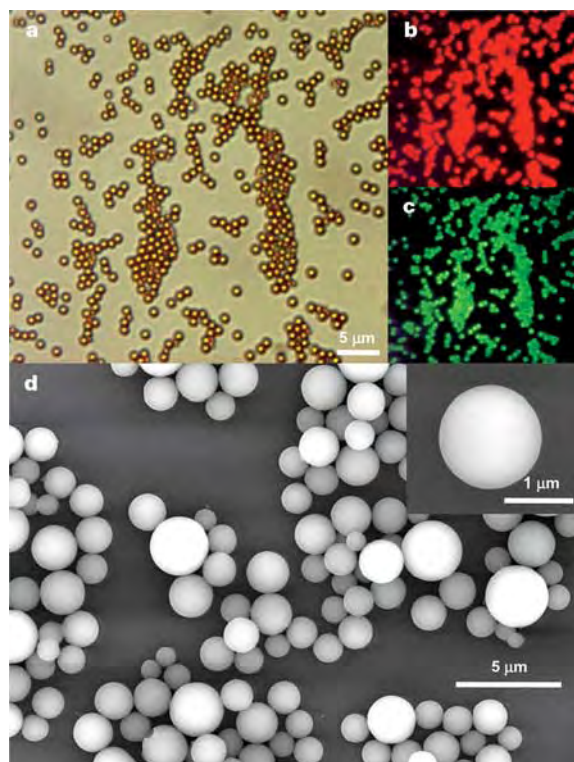


Figure 2 | Images of the spherical microparticles Zn-BMSB-Zn (3a). These have an average diameter of $1.60 \pm 0.47 \mu\text{m}$ (s.d., $n = 100$) as determined by SEM and $1.72 \mu\text{m}$ as determined by DLS. Images were obtained by optical microscopy (a), fluorescence microscopy (b, c) and SEM (d; inset in d is a high-resolution, zoom-in image).

solvothermal synthesis, layering, and slow evaporation^{18–22}. In our system, spherical microparticles Zn-BMSB-Zn, **3a** (Fig. 2), instead of the more typically observed macroscopic materials, form within one hour upon slow diffusion of diethyl ether into a precursor solution containing homochiral Zn-BMSB (**2a**) and Zn(OAc)₂ (**1a**) in a 1:1 ratio in pyridine at room temperature. The addition of diethyl ether or pentane to the polar precursor solution results in precipitation due to the low solubility of the particles in non-polar media. The resulting particles are stable in organic solvents (toluene, methanol, DMF and dimethyl sulphoxide (DMSO)), in water and in the dried state.

Optical microscopy (Fig. 2a), fluorescence microscopy (Fig. 2b and c), and scanning electron microscopy (SEM, Fig. 2d) images of example compositions show spherical particles with an average diameter of $1.60 \pm 0.47 \mu\text{m}$. Average particle size was also measured by dynamic light scattering (DLS) and was in agreement with the SEM determined value. The chemical composition of the particles was determined by energy dispersive X-ray spectroscopy and elemental analysis. Control experiments with **2d**, which has no carboxylate groups, show that the coordination polymer and, therefore, microparticles, will not form in the absence of the carboxylate groups. Infrared spectra taken before and after formation of the particles show that the carboxylate groups are coordinating to metal ions, as evidenced by a shift in CO stretching frequency from 1,653–1,659 cm^{-1} for the monomeric unbound forms (**2a–c**) to 1,597–1,605 cm^{-1} for the polymer particles (**3a–c**). In addition to the aforementioned control experiments and Fourier-transform infrared spectroscopy, we also find that ¹H and ¹³C NMR spectroscopy, and electrospray ionization mass spectrometry, are all consistent with the proposed mode of polymerization (see Supplementary Information). Featureless powder X-ray diffraction data for these particles show that they are amorphous and not crystalline materials. The

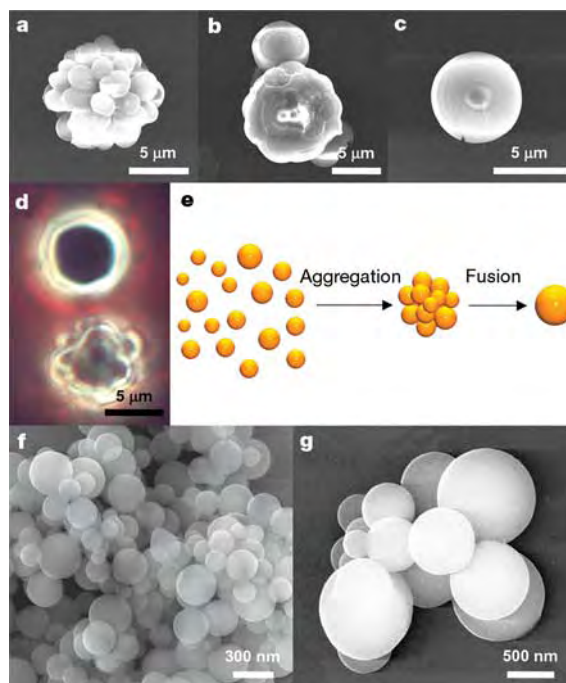


Figure 3 | SEM and optical microscopy images of example micro- and nanoparticles formed through the infinite coordination polymer strategy, and the proposed cluster-fusion growth mechanism. a–c, SEM images monitoring the growth process for Zn-BMSB-Zn (**3a**). a, Image of an early intermediate cluster formed by aggregation of small particles. The small particles become oblong as they merge with each other. b, Image of an intermediate at a later stage, showing surface annealing. c, Image of a fully formed spherical particle. d, Dark-field optical microscopy image of cluster aggregates (bottom) and a fully formed particle (top). e, A schematic representation of the proposed cluster-fusion growth mechanism. f, Image of spherical particles **3a** with an average diameter of $190 \pm 60 \text{ nm}$ (s.d., $n = 50$) as determined by SEM and 176 nm as determined by DLS. g, Image of spherical particles **3a** with an average diameter of $780 \pm 230 \text{ nm}$ (s.d., $n = 50$) as determined by SEM and 575 nm as determined by DLS.

Zn-BMSB-Zn particles (**3a**) are fluorescent in the red and green regions of the spectrum (Fig. 2b and c) as a result of the highly fluorescent BMSB building block, **2a** (ref. 24). Indeed, the emission spectra of the polymer particles **3a** and monomer building block **2a** are nearly identical (Supplementary Fig. 2).

When pentane is used as an initiation solvent instead of diethyl ether, larger spherical microparticles ($\sim 5 \mu\text{m}$) form (Fig. 3a–d, Supplementary Fig. 4). Under these synthesis conditions, the growth of the particles can be observed by taking aliquots from the synthesis mixture at various stages and characterizing the particles by SEM and optical microscopy. The observations reveal two kinds of intermediate particles: clusters formed by aggregation of small particles, and a fused version of such clusters. At early stages of the reaction, the clusters of smaller particles are readily observable (Fig. 3a and 3d, bottom); the clusters then slowly anneal into single particles with smoother surfaces (Fig. 3b) and ultimately form uniform spherical particles (Fig. 3c and 3d, top). These observations suggest a two-step cluster-fusion growth mechanism (Fig. 3e), where several small particles first aggregate to form large cluster particles, which in a second step undergo intra-particle fusion to yield large uniform spherical particles. This can occur because of the reversible nature of the metal coordination chemistry, allowing the system to anneal into a smooth particle. The cluster fusion step can involve dozens of particles or only a few, depending upon conditions, and the ultimate size of the large spherical particles depends upon the number of smaller particles involved in the fusion process. Although they cannot be completely ruled out, physical effects probably play a

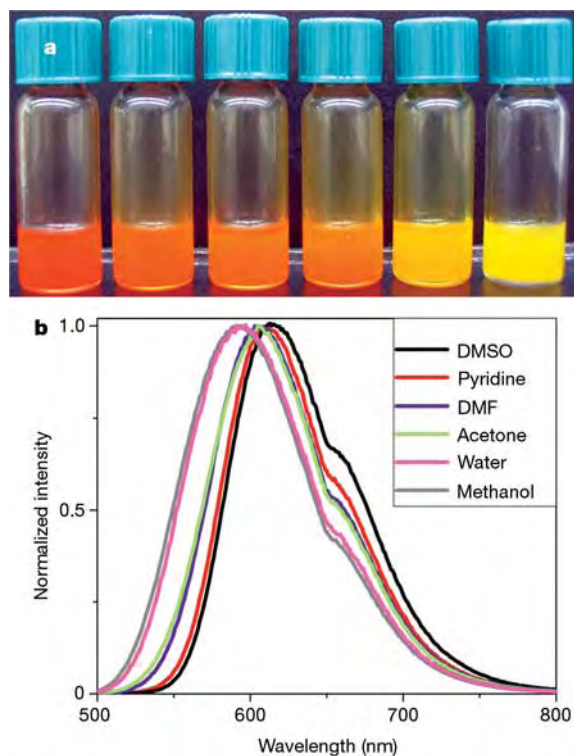


Figure 4 | Optical properties of 3a. **a**, A photograph of a series of spherical particles Zn-BMSB-Zn 3a where the ancillary ligands have been systematically changed (left to right: L = dimethyl sulphoxide (DMSO), pyridine, *N,N*-dimethylformamide (DMF), acetone, methanol and water). All particles are dispersed in toluene except where L is water, in which case pure water was used. **b**, Emission spectra of this series of particles 3a. Excitation wavelength, 420 nm.

more minor role in the polymerization process as the reaction mixture is not stirred while the particles are forming.

The rate of addition and type of initiation solvent allow control of particle size. Fast addition results in nanoscale particles by quenching the growth process at an early stage in the reaction. Nanoparticles of 3a with an average diameter of 190 ± 60 nm (Fig. 3f) could be prepared by the rapid addition of diethyl ether to the reaction mixture of 2a and 1a in pyridine. In contrast, the fast addition of pentane as the initiation solvent under nearly identical conditions resulted in particles with an average diameter of 780 ± 230 nm (Fig. 3g). The polarity of the solvent affects the solubility of the resulting particles and, therefore, their average size. This size control will probably be refined as this process evolves.

We also investigated the utility of this method for producing spherical particles of coordination polymers with other metal ions such as Cu^{2+} and Ni^{2+} (see Supplementary Information). Spherical nanoparticles of Cu-BMSB-Cu (3b), synthesized from the fast addition of pentane into a precursor solution containing Cu-BMSB (2b) and $\text{Cu}(\text{OAc})_2(\text{H}_2\text{O})$ (1b) in pyridine, have diameters similar to Zn-BMSB-Zn particles (3a) prepared via an analogous procedure. Slow diffusion of pentane into a precursor solution containing 2b and 1b yields particles that are on average significantly larger than the particles formed from the fast addition method. Similarly, the reaction between Ni-BMSB (2c) and $\text{Ni}(\text{OAc})_2 \cdot 4(\text{H}_2\text{O})$ (1c) gives spherical particles Ni-BMSB-Ni (3c). In contrast to 3a, particle compositions 3b and 3c are not fluorescent because they are not made of fluorescent BMSB building blocks.

The ancillary ligands (L in Fig. 1) in these polymer particles allow the particles' physical properties to be controlled by adjusting the electronic nature of the metals to which they are coordinated. For example, a red toluene suspension of the spherical particle

composition of 3a turns yellow when methanol is added to the reaction vessel. This colour change is attributed to the replacement of the pyridine ligand on the Zn metal centres with methanol. If the solvent is removed from this complex and the particles are redispersed in toluene with DMSO (10% by volume), the yellow solution turns red owing to the formation of Zn-DMSO adducts. It is well known that isoelectronic monomers of this class of compounds exhibit characteristic red-shifts in the solution state as a function of the increasing σ -donor properties of the ancillary ligands²⁵. The optical properties of the nanoparticles are highly dependent upon the type of coordinating ligands, and the effects of different ligands (DMSO, pyridine, DMF, acetone, methanol and water) can be observed with the naked eye (Fig. 4a). All reactions, with the exception of pyridine, which dissolves the particles (see above), appear to be completely reversible, and can be effected within seconds of introduction to the appropriate small molecule. Emission and diffuse reflectance spectra of this series of particle compositions demonstrate the reversibility of these reactions and the ability to fine-tune the optical properties of the particles through choice of ancillary ligands (Fig. 4b).

As BMSB building blocks can be prepared in enantiopure form and are important compounds in homogeneous catalysis²⁶, this new class of material may find use in asymmetric catalysis and chiral separations. In these applications, the small size of the particles results in a high overall surface area that ensures major advantages over bulk materials.

Received 8 June; accepted 19 August 2005.

- Bell, A. T. The impact of nanoscience on heterogeneous catalysis. *Science* **299**, 1688–1691 (2003).
- Wang, J., Gudiksen, M. S., Duan, X., Cui, Y. & Lieber, C. M. Highly polarized photoluminescence and photodetection from single indium phosphide nanowires. *Science* **293**, 1455–1457 (2001).
- Blanco, A. *et al.* Large-scale synthesis of a silicon photonic crystal with a complete three-dimensional bandgap near 1.5 micrometres. *Nature* **405**, 437–440 (2000).
- Cao, Y. C., Jin, R. & Mirkin, C. A. Nanoparticles with Raman spectroscopic fingerprints for DNA and RNA detection. *Science* **297**, 1536–1540 (2002).
- Bruchez, M. Jr, Moronne, M., Gin, P., Weiss, S. & Alivisatos, A. P. Semiconductor nanocrystals as fluorescent biological labels. *Science* **281**, 2013–2016 (1998).
- Taton, T. A., Mirkin, C. A. & Letsinger, R. L. Scanometric DNA array detection with nanoparticle probes. *Science* **289**, 1757–1760 (2000).
- Mirkin, C. A., Letsinger, R. L., Mucic, R. C. & Storhoff, J. J. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. *Nature* **382**, 607–609 (1996).
- Elghanian, R., Storhoff, J. J., Mucic, R. C., Letsinger, R. L. & Mirkin, C. A. Selective colorimetric detection of polynucleotides based on the distance-dependent optical properties of gold nanoparticles. *Science* **277**, 1078–1081 (1997).
- Sun, S., Murray, C. B., Weller, D., Folks, L. & Moser, A. Monodisperse FePt nanoparticles and ferromagnetic FePt nanocrystal superlattices. *Science* **287**, 1989–1992 (2000).
- Horn, D. & Rieger, J. Organic nanoparticles in the aqueous phase—theory, experiment, and use. *Angew. Chem. Int. Edn Engl.* **40**, 4331–4361 (2001).
- Roscoe, S. B., Fréchet, J. M. J., Walzer, J. F. & Dias, A. J. Polyolefin spheres from metallocenes supported on noninteracting polystyrene. *Science* **280**, 270–273 (1998).
- Fu, H.-B. & Yao, J.-N. Size effects on the optical properties of organic nanoparticles. *J. Am. Chem. Soc.* **123**, 1434–1439 (2001).
- Henglein, A. Reduction of $\text{Ag}(\text{CN})_2^-$ on silver and platinum colloidal nanoparticles. *Langmuir* **17**, 2329–2333 (2001).
- Peng, X. *et al.* Shape control of CdSe nanocrystals. *Nature* **404**, 59–61 (2000).
- Chen, J., Herricks, T. & Xia, Y. Polyol synthesis of platinum nanostructures: control of morphology through the manipulation of reduction kinetics. *Angew. Chem. Int. Edn Engl.* **44**, 2589–2592 (2005).
- Jin, R. *et al.* Photoinduced conversion of silver nanospheres to nanoprisms. *Science* **294**, 1901–1903 (2001).
- Jin, R. *et al.* Controlling anisotropic nanoparticle growth through plasmon excitation. *Nature* **425**, 487–490 (2003).
- Seo, J. S. *et al.* A homochiral metal-organic porous material for enantioselective separation and catalysis. *Nature* **404**, 982–986 (2000).
- Rosi, N. L. *et al.* Hydrogen storage in microporous metal-organic frameworks. *Science* **300**, 1127–1129 (2003).

20. Yaghi, O. M. *et al.* Reticular synthesis and the design of new materials. *Nature* **423**, 705–714 (2003).
21. Evans, O. R. & Lin, W. Crystal engineering of NLO materials based on metal-organic coordination networks. *Acc. Chem. Res.* **35**, 511–522 (2002).
22. Kosal, M. E., Chou, J.-H., Wilson, S. R. & Suslick, K. S. A functional zeolite analogue assembled from metalloporphyrins. *Nature Mater.* **1**, 118–121 (2002).
23. Tabellion, F. M., Seidel, S. R., Arif, A. M. & Stang, P. J. Template and guest effects on the self-assembly of a neutral and homochiral helix. *Angew. Chem. Int. Edn Engl.* **40**, 1529–1532 (2001).
24. Splan, K. E. *et al.* Photophysical and energy-transfer properties of (salen)zinc complexes and supramolecular assemblies. *Eur. J. Inorg. Chem.*, 2348–2351 (2003).
25. Kilså, K. *et al.* Anchoring group and auxiliary ligand effects on the binding of ruthenium complexes to nanocrystalline TiO₂ photoelectrodes. *J. Phys. Chem. B* **108**, 15640–15651 (2004).
26. Ruck, R. T. & Jacobsen, E. N. Asymmetric catalysis of hetero-ene reactions with tridentate Schiff base chromium(III) complexes. *J. Am. Chem. Soc.* **124**, 2882–2883 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements C.A.M. acknowledges the US Air Force Office of Scientific Research, NIH, NSF and DARPA for supporting this research. We thank C. L. Stern for X-ray crystallographic analysis.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.A.M. (chadnano@northwestern.edu).

Slowing of the Atlantic meridional overturning circulation at 25° N

Harry L. Bryden¹, Hannah R. Longworth¹ & Stuart A. Cunningham¹

The Atlantic meridional overturning circulation carries warm upper waters into far-northern latitudes and returns cold deep waters southward across the Equator¹. Its heat transport makes a substantial contribution to the moderate climate of maritime and continental Europe, and any slowdown in the overturning circulation would have profound implications for climate change. A transatlantic section along latitude 25° N has been used as a baseline for estimating the overturning circulation and associated heat transport^{2–4}. Here we analyse a new 25° N transatlantic section and compare it with four previous sections taken over the past five decades. The comparison suggests that the Atlantic meridional overturning circulation has slowed by about 30 per cent between 1957 and 2004. Whereas the northward transport in the Gulf Stream across 25° N has remained nearly constant, the slowing is evident both in a 50 per cent larger southward-moving mid-ocean recirculation of thermocline waters, and also in a 50 per cent decrease in the southward transport of lower North Atlantic Deep Water between 3,000 and 5,000 m in depth. In 2004, more of the northward Gulf Stream flow was recirculating back southward in the thermocline within the subtropical gyre, and less was returning southward at depth.

Some climate models suggest that the anthropogenic increase in atmospheric carbon dioxide will result in a slowdown of the Atlantic overturning circulation⁵. Coupled climate model runs that had the Atlantic overturning circulation shut off exhibited a cooling over northwest Europe with temperatures 4 °C lower than at present⁶. Thus, any indication of a slowdown in the Atlantic overturning circulation has profound implications for climate change. In March 2004 we deployed an array of moored instruments along 25° N to begin to monitor the overturning circulation⁷ and in April–May we took a transatlantic hydrographic section along 25° N to provide an initial calibration for the time-series array measurements⁸.

The 25° N transatlantic hydrographic section was occupied in 1957 (ref. 9), in 1981 (ref. 3) and again in 1992 (ref. 10). Analysis of these three occupations suggested that the overturning circulation and heat transport at 25° N had been reasonably constant with only relatively small changes in thermocline, intermediate and deep water transports^{3,4}. In 1998, the 25° N section was again occupied¹¹, so the

section of 2004 marked the fifth complete transatlantic section along 25° N. Here we analyse the new 2004 section and the 1998 section using methods similar to those previously developed for the 1957, 1981 and 1992 sections^{2,4} and examine the structure of the overturning circulation for all five sections.

Each section extends from the African continental shelf to the Bahama Islands (Fig. 1). The 1957 and 1992 sections were effectively along 24.5° N over the entire width of the Atlantic. The 1981, 1998 and 2004 sections angled southwestward from the African continental shelf at about 28° N to join the standard 24.5° N section at about 23.5° W. To take advantage of the continuous electromagnetic cable monitoring of Gulf Stream transport through the Florida Straits¹², the 1998 and 2004 sections angled northwestward at about 73° W to complete the section along 26.5° N.

The analysis calculates geostrophic velocities for each station pair along the section. A reference level of 3,200 dbar is used for station pairs east of the western boundary region where current meter observations suggest 1,000 dbar to be more suitable^{4,13}. The transition between the two reference levels is identified from the distribution of dissolved oxygen concentration that marks the eastern edge of the boundary region⁴ and ranges from 68.3° W to 70.6° W. The concept behind the analysis is to estimate the annual average overturning, so the annual averaged wind-driven surface Ekman transport and the annual averaged Gulf Stream transport through Florida Straits must be balanced by the overall southward geostrophic transport across the mid-ocean section. Thus a uniform reference level velocity is added everywhere along the section to force the mid-ocean geostrophic transport to balance the Gulf Stream plus Ekman transport. This approach assumes that the large-scale baroclinic interior flow does not vary on seasonal or shorter timescales; theoretical arguments and modelling results support such an assumption^{14,15}.

Gulf Stream transport through the Florida Straits has been reasonably constant at 32.2 Sv (1 Sv = 10⁶ m³ s⁻¹) since 1980 (refs 12, 16) with a standard deviation in annual mean transport of 1.1 Sv. Sporadic estimates of Gulf Stream transport back to the 1960s^{16–19} and cable estimates of transport since 2000 (ref. 20) show no evidence of changes in annual averaged transport through

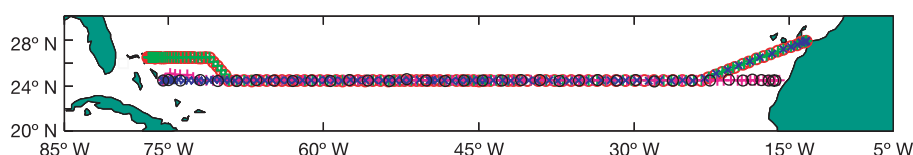


Figure 1 | Station positions for transatlantic hydrographic sections taken in 1957, 1981, 1992, 1998 and 2004. The 1957 and 1992 sections each went zonally along 24.5° N from the African coast to the Bahama Islands. Because of diplomatic clearance issues, the 1981, 1998 and 2004 sections angled

southwestward from the African coast at about 28° N to join the 24.5° N section at about 23° W. The 1998 and 2004 sections angled northwestward at about 73° W to finish the section along 26.5° N.

¹National Oceanography Centre, Empress Dock, Southampton SO14 3ZH, UK.

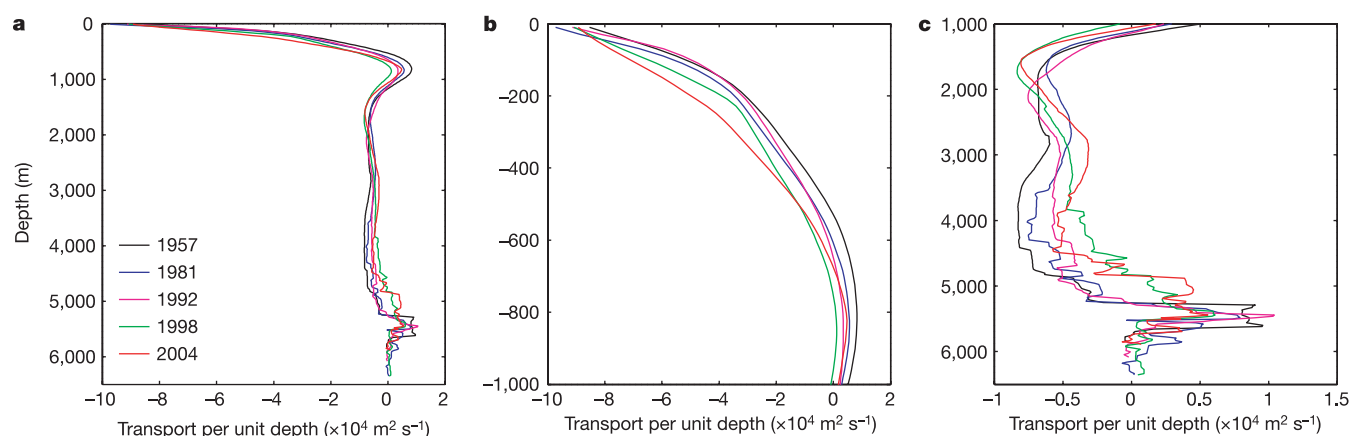


Figure 2 | Vertical distribution of mid-ocean meridional geostrophic flow across 25° N section. Transport per unit depth (in $\text{m}^2 \text{s}^{-1}$) represents the zonally averaged northward geostrophic velocity times the zonal distance across the section at each depth. **a**, Top-to-bottom profile showing the general similarity in vertical structure of the flow for each section with southward flow in the upper waters, a northward flow of intermediate

waters, a southward flow of deep waters at 1,200–5,000 m depth, and a northward flow in the bottom waters. **b**, Expanded profile of the thermocline flow showing the stronger southward flow in the 1998 and 2004 sections. **c**, Expanded profile below 1,000 m depth showing the two cores of southward flowing upper NADW centred at about 2,000 m depth and lower NADW centred at 4,000 m depth.

the Florida Straits larger than 2 Sv. In the SOC and NCEP wind stress climatologies^{21,22}, the mean Ekman transport at 25.5° N is 3.8 Sv (SOC) or 3.6 Sv (NCEP) and the variability in annual averaged Ekman transport is 0.6 Sv. There is no significant change in Ekman transport at 25° N over time in either the SOC or NCEP climatologies. There is a small net southward transport across 25° N associated with the 0.8 Sv Bering Straits throughflow from the Pacific which is diminished by a net evaporation of order 0.1 Sv over the Atlantic north of 25° N (ref. 23) but this is smaller than the uncertainty in the calculations.

Here, to be consistent with previous analyses of the 25° N sections⁴, we use a constant northward Ekman transport for each section of 5.4 Sv (ref. 24) and a constant Gulf Stream transport of 30.2 Sv for the early sections finishing at 24.5° N, and 32.2 Sv for the 1998 and 2004 sections ending at 26.5° N. The difference of 2 Sv is due to the flow through the northwest Providence channel²⁵ that joins the Gulf Stream flow north of 24.5° N to make up the 32.2 Sv measured by cable at 26.5° N. In summary, the southward mid-ocean geostrophic transport equals 35.6 Sv for the 1957, 1981 and 1992 sections and 37.6 Sv for the 1998 and 2004 sections. From the observed variability, we estimate that the uncertainty in forcing the southward mid-ocean geostrophic transport to equal a constant value for each of the five sections is only ± 2 Sv.

The overall vertical structure of the mid-ocean geostrophic circulation is similar for the five sections (Fig. 2a): there is surface-intensified southward flow in the thermocline above a depth of 800 m, small northward flow of intermediate waters between about 800 and 1,200 m, southward flow below 1,200 m down to about 5,000 m and northward flow below 5,000 m. The strength of the flows has changed, however. In the main thermocline, the southward flow is much stronger between 100 and 600 m depth in 2004 (Fig. 2b) so that the mid-ocean southward transport above 1,000 m depth has increased from 13 Sv in 1957 to nearly 23 Sv in 2004 (Table 1). In the deep waters the southward transport between 1,000 and 3,000 m that is associated with upper North Atlantic Deep Water (NADW) originating in the Labrador Sea has remained reasonably constant, varying between 9 and 12 Sv; below 3,000 m, however, the southward transport of lower NADW originating in the Greenland–Iceland Norwegian Sea has steadily decreased from 15 Sv in 1957 to 7 Sv in 2004 (Table 1). Not only has the lower-NADW transport decreased but the bottom part of the flow is gone: in 1998 and 2004 the flow passes through zero at about 4,800 m depth, whereas in earlier sections the southward flow extended down to 5,200 m (Fig. 2c).

The changes in transport distribution along 25° N are described in Supplementary Figs S1 and S2. In temperature (water mass) classes, the results are effectively the same. Thermocline waters defined to be waters warmer than 9.5 °C exhibit an increase in southward transport from 16 Sv in 1957 to 24 Sv in 2004 (Supplementary Table S1). Lower NADWs defined to have temperatures between 1.8 and 2.5 °C exhibit a consistent decrease in southward transport from 16 Sv in 1957 to 7 Sv in 1998 and 2004.

Estimates of geostrophic transport for transoceanic sections rely heavily on the stations close to the eastern and western boundaries, because these end stations effectively set the overall baroclinic shear in the currents above 1,000 m depth and the upper-level transport. The variability near the western boundary is evident in Supplementary Fig. S3, so we can argue that the upper 1,000 m transport depends critically on the nature of the western end station: for example, whether it is inside or outside an eddy. Careful consideration of the errors in geostrophic transports derived from transoceanic sections and simulated in ocean circulation models led Ganachaud to the conclusion that there is an error of ± 6 Sv in overall upper and lower layer transports²⁶. Although there is little error in overall transport owing to the constraint of basin-scale mass conservation (as discussed above), there is an uncertainty of ± 6 Sv in upper layer transport that is due to sampling in or out of eddies, and because there must be compensation by deep flows, there is also an uncertainty in deep transport of ± 6 Sv. The increased southward thermocline transport of 8 Sv and the 9 Sv decrease in lower-NADW transport in the 2004 section are close to this expected uncertainty.

Two aspects of the 2004 circulation convince us that the changes are not due to end-station variability. First, the increased southward thermocline transport is a result of substantially warmer waters in the

Table 1 | Meridional transport in depth classes across 25° N

	1957	1981	1992	1998	2004
Shallower than 1,000 m depth					
Gulf Stream and Ekman	+35.6	+35.6	+35.6	+37.6	+37.6
Mid-ocean geostrophic	−12.7	−16.9	−16.2	−21.5	−22.8
Total shallower than 1,000 m	+22.9	+18.7	+19.4	+16.1	+14.8
1,000–3,000 m	−10.5	−9.0	−10.2	−12.2	−10.4
3,000–5,000 m	−14.8	−11.8	−10.4	−6.1	−6.9
Deeper than 5,000 m	+2.4	+2.1	+1.2	+2.2	+2.5

Values of meridional transport are given in Sverdrups. Positive transports are northward.

thermocline near the Bahamas. Temperatures between 400 and 800 m depths are 1 to 2 °C warmer in 2004 and the 14 °C, 12 °C and 9.5 °C isotherms are 75 m deeper in 2004 than they were in 1957, 1981 or 1992. This warming is not restricted just to the end-stations but extends eastward from the Bahamas over several hundred kilometres. These deeper isotherms and warmer thermocline waters near the western boundary lead to a steeper slope of the thermocline across the basin and to larger overall southward geostrophic currents in the thermocline relative to a deeper reference level. The result is larger southward mid-ocean flow above 1,000 m depth and the smaller overall northward transport when the Gulf Stream and Ekman transports above 1,000 m depth are added (Table 1). Smaller net northward transport across 25° N is consonant with a reported reduction in overall northward flow through the subpolar gyre based on satellite measurements²⁷.

The second convincing aspect is that the deep compensation for the increased southward thermocline transport does not occur uniformly with depth. The simplest way for the analysis method to compensate for an 8 Sv increase in southward thermocline transport would be to adjust the entire deep-water flow uniformly so that both upper- and lower-NADW transports would decrease by about 4 Sv. Instead, the observed structure shows that the upper NADW transport from 1,000 to 3,000 m depth has not changed and effectively only the lower NADW transport has decreased. In the deep water, this change is visually evident in an upward slope of the 3 °C isotherm to the west in the 1998 and 2004 sections that was not evident in the earlier sections. This slope leads to reduced southward flow of lower NADW beneath the southward core of upper NADW at about 1,800 dbar (Fig. 2b). Thus, there have been subtle changes in the structure of the deep-water circulation, resulting only in decreased lower NADW transport. Such a reduction in lower-NADW transport is consonant with observations of the cessation of lower-NADW formation in the Norwegian–Greenland Sea²⁸ and the general freshening and weakening of the flow of lower NADW coming over the northern sills^{29,30}.

We accept that the uncertainty in transport structure for the 2004 section is ± 6 Sv in the upper and lower layer transports and that the observed changes are uncomfortably close to these uncertainties. But the warmer waters near the western boundary in the 1998 and 2004 section leading to an increase in southward mid-ocean recirculation in the thermocline and the reduction in deep water flow only in lower NADW represent strong arguments that the observed changes are robust. The decrease in net northward flow of warm upper waters and decrease in net southward flow of cold deep waters across the 25° N section result in a reduction of the northward heat transport across 25° N from 1.3–1.4 PW (1 PW = 10^{15} W) for the 1957, 1981 and 1992 sections to 1.1 PW for the 1998 and 2004 sections.

Received 8 April; accepted 28 October 2005.

1. Bryden, H. L. & Imawaki, S. in *Ocean Circulation and Climate* (eds Siedler, G., Church, J. & Gould, J.) Ch. 6.2 455–474 (Academic, London, 2001).
2. Hall, M. M. & Bryden, H. L. Direct estimates and mechanisms of ocean heat transport. *Deep-Sea Res.* **29**, 339–359 (1982).
3. Roemmich, D. & Wunsch, C. Two transatlantic sections: meridional circulation and heat flux in the subtropical North Atlantic Ocean. *Deep-Sea Res.* **32**, 619–664 (1985).
4. Lavin, A., Bryden, H. L. & Parrilla, G. Meridional transport and heat flux variations in the subtropical North Atlantic. *Glob. Atmos. Ocean Syst.* **6**, 269–293 (1998).
5. Cubasch, U. et al. in *Climate Change 2001: The Scientific Basis* (ed. Houghton, J. T.) Ch. 9, 525–582 (Cambridge Univ. Press, Cambridge, UK, 2001).
6. Vellinga, M. & Wood, R. A. Global climatic impacts of a collapse of the Atlantic thermohaline circulation. *Clim. Change* **54**, 251–267 (2002).
7. Srokosz, M. New experiment deploys observing array in N. Atlantic to investigate rapid climate change. *Eos* **85**(8), 78–83 (2004).

8. Cunningham, S. A. *RRS Discovery Cruise 279 (04 Apr–10 May 2004): A Transatlantic Hydrography Section at 24.5° N Cruise Report 54*, 1–199 (Southampton Oceanography Centre, Southampton, 2005) (<http://eprints.soton.ac.uk/17527/>).
9. Fuglister, F. C. *Atlantic Ocean Atlas of Temperature and Salinity Profiles and Data from the International Geophysical Year of 1957–1958*. Woods Hole Oceanographic Institution Atlas Series 1, 1–209 (WHOI, Woods Hole, Massachusetts, 1960).
10. Parrilla, G., Lavin, A., Bryden, H., Garcia, M. & Millard, R. Rising temperatures in the subtropical North Atlantic Ocean over the past 35 years. *Nature* **369**, 48–51 (1994).
11. Baringer, M. O'N. & Molinari, R. Atlantic Ocean baroclinic heat flux at 24 to 26° N. *Geophys. Res. Lett.* **26**, 353–356 (1999).
12. Baringer, M. O'N. & Larsen, J. C. Sixteen years of Florida Current transport at 27° N. *Geophys. Res. Lett.* **28**, 3179–3182 (2001).
13. Bryden, H. L., Johns, W. E. & Saunders, P. M. Deep western boundary current east of Abaco: Mean structure and transport. *J. Mar. Res.* **63**, 35–57 (2005).
14. Gill, A. E. & Niiler, P. P. The theory of the seasonal variability in the ocean. *Deep-Sea Res.* **20**, 141–177 (1973).
15. Jayne, S. R. & Marotzke, J. The dynamics of ocean heat transport variability. *Rev. Geophys.* **39**, 385–411 (2001).
16. Larsen, J. C. Transport and heat flux of the Florida Current at 27° N derived from cross-stream voltages and profiling data: theory and observations. *Phil. Trans. R. Soc. Lond. A* **338**, 169–236 (1992).
17. Schmitz, W. J. Jr & Richardson, W. S. On the transport of the Florida current. *Deep-Sea Res.* **15**, 679–693 (1968).
18. Richardson, W. S., Schmitz, W. J. Jr & Niiler, P. P. The velocity structure of the Florida Current from the Straits of Florida to Cape Fear. *Deep-Sea Res.* **16** (suppl.), 225–231 (1969).
19. Niiler, P. P. & Richardson, W. S. Seasonal variability of the Florida Current. *J. Mar. Res.* **31**, 144–167 (1973).
20. Meinen, C. S., Baringer, M. O. & Garcia, R. *Florida Current Transport* (<http://www.aoml.noaa.gov/phod/floridacurrent/>) (NOAA/AOML, Miami, Florida, 2005).
21. Josey, S. & Grist, J. *The NOC (formerly SOC) Air-Sea Flux Climatology* (<http://www.noc.soton.ac.uk/JRD/MET/fluxclimatology.php>) (National Oceanography Centre, Southampton, 2005).
22. Woodruff, S. *NCEP Real-time Marine Data* (<http://www.cdc.noaa.gov/cdc/data.nmc.marine.htm>) (NOAA/Climate Diagnostics Center, Boulder, Colorado, 2005).
23. Wijffels, S. E., Schmitt, R. W., Bryden, H. L. & Stigebrandt, A. Transport of freshwater by the oceans. *J. Phys. Oceanogr.* **22**, 155–162 (1992).
24. Trenberth, K. E., Large, W. G. & Olson, J. G. The mean annual cycle in global ocean wind stress. *J. Phys. Oceanogr.* **20**, 1742–1760 (1990).
25. Leaman, K. D. et al. Transport, potential vorticity, and current/temperature structure across Northwest Providence and Santaren Channels and the Florida Current off Cay Sal Bank. *J. Geophys. Res.* **100**, 8561–8569 (1995).
26. Ganachaud, A. Error budget of inverse box models: The North Atlantic. *J. Atmos. Ocean. Technol.* **20**, 1641–1655 (2003).
27. Häkkinen, S. & Rhines, P. B. Decline of subpolar North Atlantic circulation during the 1990s. *Science* **304**, 555–559 (2004).
28. Østerhus, S. & Gammelsrød, T. The abyss of the Nordic Seas is warming. *J. Clim.* **12**, 3297–3304 (1999).
29. Dickson, B. et al. Rapid freshening of the deep North Atlantic Ocean over the past four decades. *Nature* **416**, 832–837 (2002).
30. Hansen, B., Turrell, W. R. & Østerhus, S. Decreasing outflow from the Nordic seas into the Atlantic Ocean through the Faroe Bank channel since 1950. *Nature* **411**, 927–930 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The 2004 transatlantic hydrographic section along 25° N was supported by the Natural Environment Research Council as part of the Core Strategic Research Programme 'Ocean Variability and Climate' at Southampton Oceanography Centre. Analysis of the five sections along 25° N was also supported by NERC as part of the Rapid Programme. Comments on an earlier draft by J. Hirschi, W. Johns, S. Josey, C. Meinen, G. Parrilla, P. Rhines, P. Saunders, J. Toole, P. Vélez and R. Wood led to substantial improvement.

Author Contributions All authors contributed equally to this work.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.L.B. (h.bryden@noc.soton.ac.uk).

LETTERS

Density dependence explains tree species abundance and diversity in tropical forests

Igor Volkov¹, Jayanth R. Banavar¹, Fangliang He², Stephen P. Hubbell^{3,4} & Amos Maritan⁵

The recurrent patterns in the commonness and rarity of species in ecological communities—the relative species abundance—have puzzled ecologists for more than half a century^{1,2}. Here we show that the framework of the current neutral theory in ecology^{3–10} can easily be generalized to incorporate symmetric density dependence^{11–14}. We can calculate precisely the strength of the rare-species advantage that is needed to explain a given RSA distribution. Previously, we demonstrated that a mechanism of dispersal limitation also fits RSA data well^{3,4}. Here we compare fits of the dispersal and density-dependence mechanisms for empirical RSA data on tree species in six New and Old World tropical forests and show that both mechanisms offer sufficient and independent explanations. We suggest that RSA data cannot by themselves be used to discriminate among these explanations of RSA patterns¹⁵—empirical studies will be required to determine whether RSA patterns are due to one or the other mechanism, or to some combination of both.

Ecologists have long sought to explain the high levels of tree diversity that often occur in tropical forests. One aspect of this challenge is to understand the evolutionary origin and maintenance of this diversity on large spatial and temporal scales¹⁶. Another is to understand how such extraordinarily high alpha (local) tree diversity can be maintained on very local scales in particular tropical forests. For example, there are over a thousand tree species in a 52-hectare plot in Borneo (Lambir, Sarawak, Table 1). Numerous mechanisms have been proposed to explain tropical tree species coexistence on local scales; many of these hypotheses invoke density- and frequency-dependent mechanisms. Two of the most prominent of these hypotheses are the Janzen–Connell hypothesis^{11,12} and the Chesson–Warner hypothesis¹³.

The Janzen–Connell hypothesis is that seeds that disperse farther away from the maternal parent are more likely to escape mortality from host-specific predators or pathogens. This spatially structured mortality disfavors the population growth of locally abundant species relative to uncommon species by reducing the probability of species' self-replacement in the same location in the next generation.

The Chesson–Warner hypothesis is that a rare-species reproductive advantage arises when species have similar per capita rates of mortality but reproduce asynchronously, and there are overlapping generations. Processes that hold the abundance of a common species in check inevitably lead to rare-species advantage because the space or resources freed up by density-dependent deaths are then exploited by less-common species. Therefore, among-species frequency dependence is the community-level consequence of within-species density dependence, and thus they are two different manifestations of the same phenomenon. There is accumulating empirical evidence that such density- and frequency-dependent

processes may play a large part in maintaining the diversity of tropical tree communities^{17–22}.

Density and frequency dependence are familiar mechanisms in population biology, but it is surprising how rarely their consequences for species diversity and RSA in communities have been discussed (see Ch. 3 of ref. 4). Here we show that these mechanisms are sufficient to explain precisely the species abundance patterns in six tropical forest communities on three continents.

The neutral theory of biodiversity provides a convenient theoretical framework for linking community diversity patterns to the fundamental mechanisms of population biology (such as birth, death and migration) and speciation⁴. The celebrated statistical distribution for RSA, Fisher's log-series¹, can be shown to arise directly from the stochastic equations of population growth under neutrality at the speciation-extinction equilibrium. More significantly, Fisher's log-series arises when the birth and death rates are density-independent.³

According to the theory, the mean number of species with n individuals, $\langle \phi_n \rangle$, in a community at the stochastic speciation–extinction equilibrium takes the general form:

$$\langle \phi_n \rangle = SP_0 \left\langle \prod_{i=0}^{n-1} \frac{b_{i,k}}{d_{i+1,k}} \right\rangle_k$$

where $\langle \dots \rangle_k$ represents the arithmetic average over all species, S is the average number of species present in the ecosystem, P_0 is a constant, and $b_{i,k}$ and $d_{i,k}$ are birth and death rates for the k th species with i individuals. Here we have subscripted the birth and the death rates for arbitrary species k to indicate that these rates could, in principle, be species-specific for an asymmetric community. In contrast, “symmetry occurs at the species level when no change in community dynamics or the fates of individuals occurs upon switching the species of any two given populations in the community. Any given population behaves as it would previously, despite its new species label, and its effects on other populations remain the same, regardless of their species labels.” (P. Chesson, personal communication).

We note that what is important in determining the mean number of species, $\langle \phi_n \rangle$, are not the absolute rates of birth or death but their ratio, $r_{i,k} = \frac{b_{i,k}}{d_{i+1,k}}$. Indeed, $\langle \phi_n \rangle$ is proportional to $\langle r_{1,k} r_{2,k} \dots r_{n-1,k} \rangle_k$. This formulation is sufficiently general to represent communities of either symmetric or asymmetric species. Such a situation could arise, for example, from niche differences or from differing immigration fluxes resulting from the different relative abundances of the species in the metacommunity. Hereafter, however, we consider only the symmetric case of a community of non-interacting species with identical vital demographic rates. For large community size, this formulation is equivalent to the case of zero-sum dynamics studied by Hubbell⁴ (see Supplementary Information).

¹Department of Physics, 104 Davey Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802, USA. ²Department of Renewable Resources, University of Alberta, Edmonton, AB T6G 2H1, Canada. ³Department of Plant Biology, The University of Georgia, Athens, Georgia 30602, USA. ⁴The Smithsonian Tropical Research Institute, Box 2072, Balboa, Panama. ⁵Dipartimento di Fisica ‘G. Galilei’, Università di Padova, via Marzolo 8, 35131 Padova, Italy.

Table 1 | Model parameters for the six data sets

Dispersal limitation model	S	J	θ_1	m	θ_2	c	x
BCI, Panama	225	21457	48.1	0.09	47.5	1.80	0.9978
Yasuni, Ecuador	821	17546	204.2	0.43	213.2	0.51	0.9883
Pasoh, Malaysia	678	26554	192.5	0.09	189.5	1.95	0.9932
Korup, Cameroon	308	24591	52.9	0.54	53.0	0.24	0.9979
Lambir, Malaysia	1004	33175	288.8	0.11	301.0	2.02	0.9915
Sinharaja, Sri Lanka	167	16936	27.3	0.55	28.3	0.38	0.9983

Model comparison	L_1	L_2	Deviance	P -value
BCI, Panama	-314.0	-315.0	2.0	0.16
Yasuni, Ecuador	-301.0	-303.6	5.2	0.02
Pasoh, Malaysia	-363.7	-365.3	3.2	0.07
Korup, Cameroon	-322.3	-323.1	1.6	0.21
Lambir, Malaysia	-390.5	-391.2	1.4	0.24
Sinharaja, Sri Lanka	-258.9	-258.5	0.8	0.37

Maximum-likelihood estimates of the dispersal limitation model³ and the density-dependent symmetric model parameters (upper table) and comparison between the models (lower table) for the six data sets of tropical forests. In the six plots coordinated by Center for Tropical Forest Science of the Smithsonian (<http://www.ctfs.si.edu>), we considered trees with diameter at breast height ≥ 10 cm. S is the number of species, J is the total abundance and θ_1 and θ_2 are the biodiversity parameters in the dispersal limitation model³ and equation (1) respectively (note that θ_2 is a function of c , x and S and that both models have the same number of fitting parameters). The comparison of the models was carried out with the likelihood ratio test^{10,23,24}. The lower table presents deviance (twice the difference in the log-likelihoods L_1 and L_2 of the dispersal limitation model³ and the density-dependent symmetric model respectively) between the two models and the corresponding P -value of the χ^2 -distribution with one degree of freedom. The main result is that the dispersal-limitation model and the simple symmetric density-dependent model presented here are statistically comparable to each other in their ability to fit the tropical forest data.

We define $\hat{r}_n = \frac{b_n}{d_{n+1}} \frac{n+1}{n}$, where the factor $\frac{n+1}{n}$ is chosen to obtain $\hat{r}_n = x$ for the Fisher log-series $\langle \phi_n \rangle \propto x^n/n$. In Fisher's case, $\hat{r}_n$ does not change with population density and is an intraspecific parameter that measures the relative vital rates of birth and death of a population. To obtain intraspecific density dependence, $\hat{r}_n$ becomes a function of the population density n . Within our framework, $\frac{\langle \phi_{n+1} \rangle}{\langle \phi_n \rangle} = \frac{n}{n+1} \hat{r}_n$.

We now introduce the modified symmetric theory that captures density dependence (rare-species advantage or common-species disadvantage). In the modified theory, $\hat{r}_n$ will be a decreasing function of abundance, thereby incorporating density dependence. The equations of density dependence in the per capita birth and death rates for an arbitrary species of abundance n are:

$$\frac{b(n)}{n} = b \times \left[1 + \frac{b_1}{n} + o\left(\frac{1}{n^2}\right) \right]$$

and

$$\frac{d(n)}{n} = d \times \left[1 + \frac{d_1}{n} + o\left(\frac{1}{n^2}\right) \right]$$

for $n > 0$ as the leading terms of a power series in $(1/n)$, $\frac{b(n)}{n} = b \times \sum_{l=0}^{\infty} b_l n^{-l}$ and $\frac{d(n)}{n} = d \times \sum_{l=0}^{\infty} d_l n^{-l}$, where b_l and d_l are constants. This expansion captures the essence of density dependence by ensuring that the per capita birth rate to death rate ratios decrease and approach a constant value for large n . This happens because the higher-order terms are negligible. Note that the quantity that controls the RSA distribution is the ratio b_n/d_{n+1} . Thus the birth and death rates, b_n and d_n , are defined up to multiplicative factors $f(n+1)$ and $f(n)$ respectively, where f is any arbitrary well-behaved function.

We expect that the per capita birth rate or the fecundity will go down as the abundance increases, whereas the mortality ought to increase with abundance. Indeed, the per capita death rate can be arranged to be an increasing function of n , as observed in nature, by choosing an appropriate function f and adjusting the birth rate appropriately so that the ratio b_n/d_{n+1} remains the same. For example, the choice $f(n) = n/(n+c)$ yields a constant per capita death rate $d_n = dn$ and a fecundity that decreases with increasing abundance.

This mathematical formulation of density dependence may seem unusual to ecologists familiar with the logistic or Lotka-Volterra systems of equations, in which density dependence is typically described as a polynomial expansion of powers of n truncated at the quadratic level. However, this classical expansion is not valid in

our context because the range of n is from 1 to an arbitrarily large value, not to some fixed carrying capacity. Therefore an expansion in terms of powers of $(1/n)$ is more appropriate. For this symmetric model, noting that $\langle \phi_n \rangle = SP_0 \prod_{i=0}^{n-1} \frac{b_i}{d_{i+1}}$, we readily arrive at the following relative species-abundance relationship:

$$\langle \phi_n \rangle = \theta \frac{x^n}{n+c} \quad (1)$$

where $x = b/d$, and for parsimony we have made the simple assumption that $b_1 = d_1 = c$. The biodiversity parameter θ is the normalization constant that ensures that the average number of species in the community is S and is given by $\theta = S \frac{1+c}{c} F^{-1}(1+c, 2+c, x)$, where $F(1+c, 2+c, x)$ is the standard hypergeometric function. The parameter c measures the strength of the symmetric density dependence in the community, and it controls the shape of the RSA distribution. Note that when $c \rightarrow 0$ (the case of no density dependence), we obtain the Fisher log-series. In this case, as shown in ref. 3, θ captures the effects of speciation.

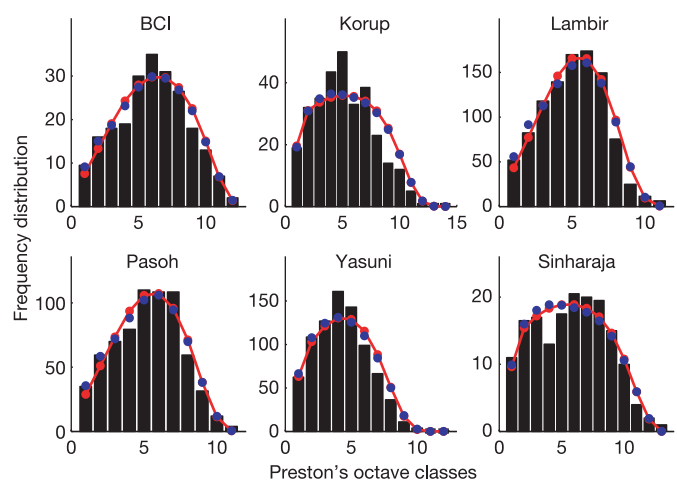


Figure 1 | Fits of density-dependent symmetric model (red line) and dispersal-limitation model³ (blue circles) to the tree species abundance data from the BCI, Yasuni, Pasoh, Lambir, Korup and Sinharaja plots, for trees ≥ 10 cm in stem diameter at breast height (see Table 1). The frequency distributions are plotted using Preston's binning method as described in ref. 3. The numbers on the x axis represent Preston's octave classes.

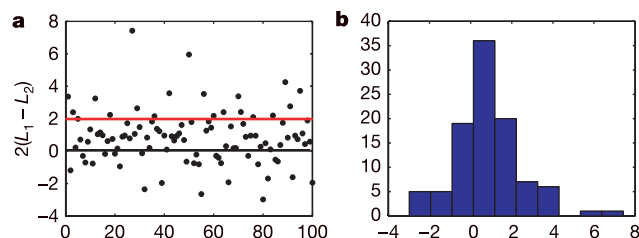


Figure 2 | Test of the equivalence of the dispersal limitation model and the density-dependent symmetric model. We have randomly generated 100 communities using the dispersal-limitation model, following the Etienne²³ algorithm with the values of m and θ_1 given in Table 1 for the BCI plot ($m = 0.09$ and $\theta_1 = 48.1$). For each of these data sets we calculated the log-likelihoods L_1 and L_2 of the dispersal-limitation model³ and the density-dependent symmetric model respectively. **a**, The deviances for each of the data set; **b**, the corresponding histogram. The red line in **a** corresponds to the deviance estimated from the BCI data. The deviance is distributed almost symmetrically around zero (with a mean value of 0.85) and the BCI data is indistinguishable from the random pseudo-samples.

We now show how equation (1) estimates the strength of symmetric density dependence that is consistent with the observed RSA distributions of tree species in six large tropical forest plots on three continents: Barro Colorado Island (BCI), Panama; the Yasuni National Park, Ecuador; the Pasoh Forest Reserve, peninsular Malaysia; the Korup National Park, Cameroon; the Lambir Hills National Park, Sarawak, Malaysia; and the Sinharaja World Heritage Site, Sri Lanka. These site plots are part of a global network of large plots managed by the Center for Tropical Forest Science of the Smithsonian Tropical Research Institute. These New and Old World tropical forests have had long separate ecological and evolutionary histories, but despite these different histories, the symmetric theory with density dependence fits each of the RSA distributions very well.

Figure 1 shows the fits of equation (1) and the dispersal limitation model³ to the tree abundance data collected from the six permanent plots of tropical forest. These plots are 50 hectares except for Lambir (52 hectares), Yasuni (25 hectares) and Sinharaja (25 hectares). The results in Table 1 and Fig. 1 show that the RSA data of tree species in these plots are equally well described both by the density-dependent model and the dispersal-limitation model³ (also see Fig. 2).

The rare-species advantage is illustrated in Fig. 3 and is of the same order of magnitude in the different forests. The key quantity that controls the RSA is the birth rate to death rate ratio $\hat{r}_n$ defined above. The curves in Fig. 3 were derived from the parameters in Table 1, which in turn were obtained from the empirical RSA data using the maximum-likelihood method. At stochastic steady state, community size (mass balance) is maintained by the slow rate of decline in common species (at large n in Fig. 3) exactly balanced by the growth of rare species, and by the very slow input of new species by speciation.

Several important ecological insights result from this new theory. First, we have shown that an assumption of asymmetric density dependence, for example, postulating different carrying capacities for each species, is not necessary to explain patterns of RSA at least in these six tropical forests: a much simpler symmetric hypothesis is sufficient. Second, we have shown that the population sizes that exhibit rare-species advantage consistent with the observed RSA data are all quite small. The transition to a 'Fisher log-series'-like value for $x = b/d$ that is slightly less than replacement occurs at what would be considered low population densities of tree species in these forests (<1 tree per hectare). Third, applied to spatially mapped ecological communities, the theory reveals the scales on which density dependence occurs and on what larger scales it gives way to density independence. In these six tropical forests, the theory is consistent with density dependence operating at very similar and small scales of abundance and spatial length, and weakening at larger scales.

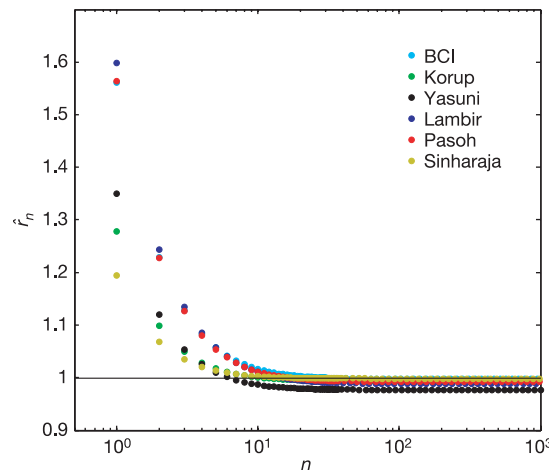


Figure 3 | Plot of $\hat{r}_n$ derived from equation (1) versus n for the six data sets of tropical trees. For large values of n , $\hat{r}_n$ asymptotes to a value slightly less than 1. The BCI data (cyan circles) at small n is almost invisible because it coincides with the Pasoh data set (red circles).

Finally, we have demonstrated that symmetric density dependence gives an equally sufficient mechanistic explanation for RSA patterns, in addition to and independent of dispersal limitation³. In Table 1, we show the fits of the two mechanisms to the RSA data from the six forests, from which it is clear that both mechanisms yield fits that cannot be distinguished statistically in quality. However, the ecological explanation that accompanies each of these mechanisms is very different. According to the dispersal mechanism, the explanation for the lower frequency of rare species compared to species of middling abundance is that rare species are more extinction-prone, and when they go extinct in a community, they take longer to re-immigrate than common species do. According to the density-dependence mechanism, on the other hand, the reduced steady-state frequency of rare species arises because populations of rare species grow differentially faster into higher abundance categories owing to a rare-species advantage. An important conclusion is that we cannot deduce the mechanisms causing a particular RSA pattern from RSA data alone (see Supplementary Information). Because these mechanisms are not mutually exclusive, it must be left to empirical research to uncover the relative contributions of each mechanism to observed RSA patterns. However, we do note one distinction between the two mechanisms. The dispersal-limitation mechanism generally implies that we are considering a local community into which immigration is possible. However, the density dependence mechanism can apply equally well to local communities or to the metacommunity.

Received 13 May; accepted 7 July 2005.

1. Fisher, R. A., Corbet, A. S. & Williams, C. B. The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12**, 42–58 (1943).
2. Preston, F. W. The commonness and rarity of species. *Ecology* **29**, 254–283 (1948).
3. Volkov, I., Banavar, J. R., Hubbell, S. P. & Maritan, A. Neutral theory and relative species abundance in ecology. *Nature* **424**, 1035–1037 (2003).
4. Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. New Jersey, 2001).
5. Bell, G. The distribution of abundance in neutral communities. *Am. Nat.* **155**, 606–617 (2000).
6. McKane, A., Alonso, D. & Solé, R. V. Mean-field stochastic theory for species-rich assembled communities. *Phys. Rev. E* **62**, 8466–8484 (2000).
7. Bell, G. Neutral macroecology. *Science* **293**, 2413–2418 (2001).
8. Vallade, M. & Houchmandzadeh, B. Analytical solution of a neutral model of biodiversity. *Phys. Rev. E* **68**, 061902 (2003).
9. Houchmandzadeh, B. & Vallade, M. Clustering in neutral ecology. *Phys. Rev. E* **68**, 061912 (2003).

10. Alonso, D. & McKane, A. J. Sampling Hubbell's neutral theory of biodiversity. *Ecol. Lett.* **7**, 911–914 (2004).
11. Janzen, D. H. Herbivores and the number of tree species in tropical forest. *Am. Nat.* **104**, 501–528 (1970).
12. Connell, J. H. in *Dynamics of Populations* (eds Den Boer, P. J. & Gradwell, G. R.) 298–312 (Centre for Agricultural Publishing and Documentation, Wageningen, The Netherlands, 1971).
13. Chesson, P. L. & Warner, R. R. Environmental variability promotes coexistence in lottery competitive systems. *Am. Nat.* **117**, 923–943 (1981).
14. Chesson, P. A need for niches? *Trends Ecol. Evol.* **6**, 26–28 (1991).
15. Etienne, R. S. & Olff, H. A novel genealogical approach to neutral biodiversity theory. *Ecol. Lett.* **7**, 170–175 (2004).
16. Morley, R. J. *Origin and Evolution of Tropical Rainforests* Ch. 5 (Wiley, New York, 2000).
17. Augspurger, C. K. Seedling survival of tropical tree species: interactions of dispersal distance, light gaps, and pathogens. *Ecology* **65**, 1705–1712 (1984).
18. Hubbell, S. P., Condit, R. & Foster, R. B. Presence and absence of density dependence in a neotropical tree community. *Trans. R. Soc. Lond.* **330**, 269–281 (1990).
19. Gilbert, G. S., Hubbell, S. P. & Foster, R. B. Density and distance-to-adult effects of a canker disease of trees in a moist tropical forest. *Oecologia* **98**, 100–108 (1994).
20. Condit, R., Hubbell, S. P. & Foster, R. B. Density dependence in two understory tree species in a neotropical forest. *Ecology* **75**, 671–680 (1994).
21. Harms, K. E., Wright, S. L., Calderón, O., Hernández, A. & Herre, E. A. Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature* **404**, 493–495 (2000).
22. Wright, S. J. Plant diversity in tropical forests: a review of mechanisms of species coexistence. *Oecologia* **130**, 1–14 (2002).
23. Etienne, R. S. A new sampling formula for neutral biodiversity. *Ecol. Lett.* **8**, 253–260 (2005).
24. Hilborn, R. & Mangel, M. *The Ecological Detective: Confronting Models with Data* Ch. 7 (Princeton Univ. Press, Princeton, New Jersey, 1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are indebted to D. Alonso and P. Chesson for advice. We gratefully acknowledge the work of the principal investigators and their field assistants for collecting the field data on the large plots of tropical forest. Specifically, we thank D. Thomas, G. Chuyong and D. Kenfack for the data from Korup National Park, Cameroon; R. Valencia, R. Foster and R. Condit for the data from Yasuni National Park, Ecuador; S. Davies, S. Tan, J. LaFrankie and P. Ashton for the data from Lambir Hills National Park, Sarawak; M. N. Supardi, P. Ashton and J. LaFrankie for the data from Pasoh Forest Reserve, peninsular Malaysia; and S.P.H.'s collaborators on the Barro Colorado Island plot, R. Foster and R. Condit. We also thank E. Losos for directing and coordinating the global programmes of the Center for Tropical Forest Science, which manages the plots, S. Loo for data management, and I. Rubinoff. This work was supported by NASA, by the NSF and the NSERC (Canada). The fieldwork has also received long-term support from the John D. and Catherine T. MacArthur Foundation, the Mellon Foundation, Earthwatch, Frank Levinson and the Celera Foundation, and other private foundations and individual donors.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.R.B. (banavar@psu.edu) or A.M. (maritan@pd.infn.it).

LETTERS

Glyoxalase 1 and glutathione reductase 1 regulate anxiety in mice

Iris Hovatta¹, Richard S. Tennant¹, Robert Helton^{1,2}, Robert A. Marr¹, Oded Singer¹, Jeffrey M. Redwine³, Julie A. Ellison¹, Eric E. Schadt⁴, Inder M. Verma¹, David J. Lockhart¹ & Carrolee Barlow^{1,2}

Anxiety and fear are normal emotional responses to threatening situations. In human anxiety disorders—such as panic disorder, obsessive-compulsive disorder, post-traumatic stress disorder, social phobia, specific phobias and generalized anxiety disorder—these responses are exaggerated. The molecular mechanisms involved in the regulation of normal and pathological anxiety are mostly unknown. However, the availability of different inbred strains of mice offers an excellent model system in which to study the genetics of certain behavioural phenotypes^{1–3}. Here we report, using a combination of behavioural analysis of six inbred mouse strains with quantitative gene expression profiling of several brain regions, the identification of 17 genes with expression patterns that correlate with anxiety-like behavioural phenotypes. To determine if two of the genes, glyoxalase 1 and glutathione reductase 1, have a causal role in the genesis of anxiety, we performed genetic manipulation using lentivirus-mediated gene transfer. Local overexpression of these genes in the mouse brain resulted in increased anxiety-like behaviour, while local inhibition of glyoxalase 1 expression by RNA interference decreased the anxiety-like behaviour. Both of these genes are involved in oxidative stress metabolism, linking this pathway with anxiety-related behaviour.

Different inbred mouse strains have different physical and behavioural phenotypes that are heritable and stable^{1–3}. We combined gene expression profiling and behavioural testing of multiple highly characterized strains in search of candidate genes for anxiety-like behaviour. We identified several strong candidates and performed follow-up functional studies to demonstrate directly that altered expression levels of the identified genes affected anxiety-like behaviour in mice (Supplementary Fig. 1).

Several methods to test levels of anxiety-like behaviour in mice have been developed and pharmacologically ‘validated’; that is, shown to be specifically responsive to agents with proven anxiolytic or anxiogenic effects⁴. We used two such tests to measure anxiety-like behaviour in six inbred mouse strains—the light–dark box test and the open-field test (described in the Supplementary Methods). Strain characterization with both tests was consistent (Pearson coefficient of correlation between the ‘open-field time spent in the middle of the chamber’ and the ‘light–dark box time spent in the light compartment’ was high, $r = 0.84$), and showed that A/J, DBA/2J and 129S6/SvEvTac were the most anxious strains and FVB/NJ the least anxious strain (Fig. 1a), as reported previously^{5,6}. The behaviour of C3H/HeJ and C57BL6/J animals was intermediate (Fig. 1a). In contrast, although not completely ruling out an association between locomotor activity and anxiety-like behaviour, the strain order for locomotor activity, estimated as the distance travelled in the dark compartment of the light–dark box, was different from the strain

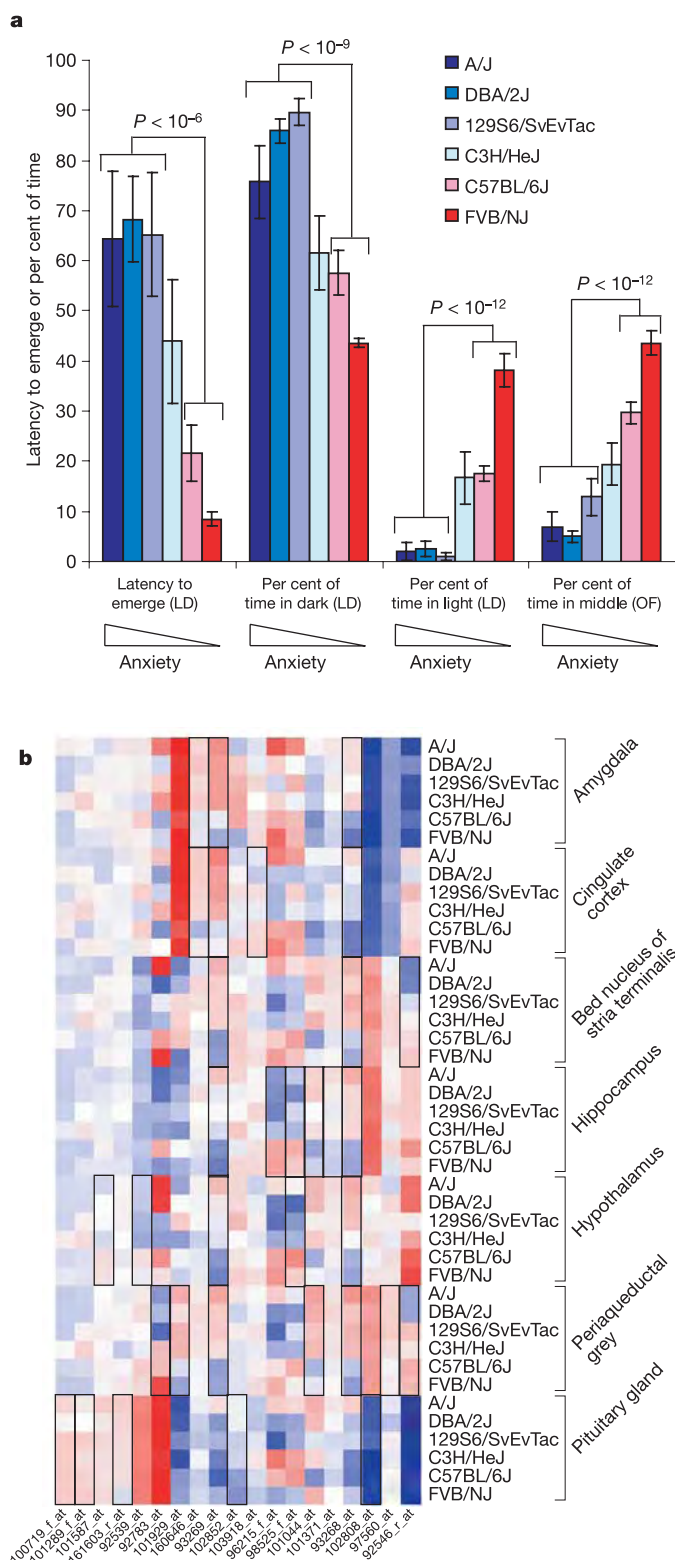
order for anxiety-like behaviour (Supplementary Information).

Several methods have been used to show that the amygdala, septohippocampal system, medial hypothalamus, central periaqueductal grey, and frontal and cingulate cortices are important brain structures involved in the regulation of anxiety and fear^{7–10}. Based on this information, we selected seven brain regions (the amygdala, bed nucleus of the stria terminalis, cingulate cortex, hippocampus, hypothalamus, periaqueductal grey and pituitary gland) thought to regulate aspects of anxiety-related behaviour, and used oligonucleotide arrays (Affymetrix U74Av2) to assess the expression levels of ~10,000 genes in those regions. To ensure that our experimental methodology and data analysis methods minimized the number of false positives and maximized the reliability of the results, we carefully compared at least two independent replicate samples for each brain region from each strain¹¹. Reproducibility between replicates was high (Supplementary Table 1), and the estimated false positive rate was low (0.013%; see the Supplementary Methods for details).

We identified oligonucleotide probe sets that showed statistically significant differences in expression levels between two of the most anxious (A/J and DBA/2J) and the two least anxious (FVB/NJ and C57BL/6J) mouse strains in at least one brain region (see the Supplementary Methods for details). We identified eight probe sets in the hippocampus, 12 in hypothalamus, 33 in pituitary, seven in bed nucleus of the stria terminalis, 19 in periaqueductal grey, 12 in amygdala and 12 in cingulate cortex. These probe sets cover genes that are differentially expressed between the phenotypic extremes, but may not necessarily correlate with anxiety-like phenotypes across all six inbred strains. Therefore, we performed a correlation analysis to identify a subset of genes with expression levels that correlate with anxiety-related phenotypes across all strains (see the Supplementary Methods for details). Nineteen probe sets were identified (Table 1, Fig. 1b and Supplementary Table 2), corresponding to 17 candidate genes (probe sets 93268_at and 93269_at both represented glyoxalase 1 (*Glo1*), and probe sets 96215_f_at and 98525_f_at both represented erythroid differentiation regulator 1 (*Erdr1*)). In addition to the correlation analysis described above, we analysed the data with a standard implementation of a linear mixed-effects model to assess the correlation between expression and anxiety-related behaviour (Table 1 and Supplementary Table 2). Only growth hormone (probe set 92783_at) did not show a statistically significant association using this method. Some of the identified genes showed differential expression across several brain regions, while the majority of the genes were differentially expressed between strains in only a single brain region (Table 1). To independently confirm the differences, we performed quantitative polymerase chain reaction with reverse transcription (quantitative RT–PCR; qPCR) for 11 of the 17

¹The Salk Institute for Biological Studies, Laboratory of Genetics, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. ²BrainCells Inc., 10835 Road to the Cure, San Diego, California 92121, USA. ³Neurome Inc., 11149 North Torrey Pines Road, La Jolla, California 92037, USA. ⁴Rosetta Inpharmatics LLC, Merck & Co., 12040 115th Avenue Northeast, Seattle, Washington 98109, USA.

candidate genes (Supplementary Fig. 2). For most of the genes, the differences in gene expression observed by microarray analysis were confirmed by qPCR. Two genes—cadherin 2 (*Cdh2*) and epoxide hydrolase 1 (*Ephx1*)—did not show clear differential expression between the strains by qPCR. It is possible that not all of these differentially expressed genes are involved with the regulation of anxiety. For example, some of them might correlate with the phenotype by chance, so we addressed this question using functional and genetic studies.



Notably, five of the 17 candidate genes were enzymes. Enzyme activity assays were available for three of them. We measured the activities of delta-aminolevulinic acid dehydratase (Alad), glyoxalase 1 (Glo1) and glutathione reductase 1 (Gsr) from brain homogenates containing combined tissue of hippocampus, striatum and cortex (Supplementary Fig. 2). It seemed that *Alad* mRNA levels in FVB/NJ animals were overestimated by the microarrays, as *Alad* expression and Alad activity did not correlate with anxiety-like behaviour across the strains. In contrast, both Glo1 and Gsr enzyme activities matched the pattern found in both the microarray and qPCR analyses, with highest activities in the most anxious and lowest activities in the least anxious strains. This was particularly intriguing given that reduced glutathione (GSH), the levels of which are maintained by Gsr, is a major antioxidant in the brain. Glo1 uses GSH as a cofactor to detoxify cytotoxic methylglyoxal. Furthermore, erythrocytes from patients with anxiety disorders (such as panic disorder or obsessive-compulsive disorder) may have higher levels of antioxidant enzymes (glutathione peroxidase and superoxide dismutase)^{12,13}, suggesting that free radicals may have a role in the pathogenesis of anxiety disorders. Oxidative stress has also been implicated in the pathogenesis of other neuropsychiatric diseases, including schizophrenia and major depressive disorder^{14,15}, and Glo1 is linked to diabetes¹⁶, Alzheimer's disease¹⁷, autism¹⁸ and the regulation of theta oscillations during sleep¹⁹. A recent study suggested Glo1 might be a biological marker for trait anxiety in bidirectionally crossed mouse lines²⁰. Therefore, we sought to determine the role of these candidate genes in influencing anxiety-related behaviour in a complex genetic background.

We analysed the offspring of two different F_1 crosses of the non-anxious C57BL/6J strain and an anxious A/J strain (AB6 F_1 and B6AF $_1$), in addition to BALB/cByJ inbred mice as this strain was shown to be very anxious. In both open-field and light-dark box tests, F_1 animals derived from the A/J and C57BL/6J crosses showed intermediate levels of anxiety-like behaviour compared to the parental strains (Fig. 2a). We hypothesized that if Glo1 and Gsr exert a strong influence on the phenotype, the activity levels of the enzymes should correlate with the anxiety-related phenotype. As expected, there was a statistically significant correlation between the open-field behaviour and the Glo1 ($P = 0.0005$) and Gsr ($P = 0.009$) enzyme activities, as measured by regression analysis over A/J, C57BL/6J, their F_1 offspring and BALB/cByJ mice (Fig. 2b and c), suggesting that these two enzymes are very strong candidates for regulating anxiety-related behaviours.

To further investigate the role of *Glo1* and *Gsr* in anxiety, we prepared lentiviral vectors to overexpress these genes *in vivo* (Supplementary Fig. 3a). The lentiviral approach was favoured over other viral vectors because lentiviral vectors efficiently transduce central nervous system cells and are not cytotoxic^{21,22}. One microlitre of either *Glo1*- or *Gsr*-containing virus, or a green fluorescent protein (GFP)-containing control virus, was injected bilaterally in the region

Figure 1 | Inbred mouse strains have different levels of anxiety-related behaviours. **a**, Behavioural tests on inbred strains of mice. Test parameters are shown on the x axis. The y axis shows either the latency to emerge from the dark side to the light side of the light-dark (LD) chamber (zero corresponds to 0 min and 100 corresponds to 5 min), the per cent of time in the dark or light side of the light-dark chamber, or the per cent of time in the middle of the open-field (OF) chamber. See the Supplementary Methods for the test measures and analysis. Values are mean $\pm$ s.e.m. P values calculated using a two-tailed Student's *t*-test. **b**, A heat map based on the cluster analysis of the 19 probe sets with signals that correlated with the anxiety-related phenotype, and that were significantly different between the most and the least anxious strains (bordered by a black box). The x axis shows the probe set identifiers. Mouse strains are organized by tissue and level of anxiety-like behaviour on the y axis. Red represents high and blue represents low signal intensity, with a more intense colour showing relatively higher signal intensity.

Table 1 | Correlation of gene expression patterns with anxiety-related phenotypes in six inbred mouse strains

Probe set	Gene title	Gene symbol	Tissue	Average fold change*	Correlation coefficient (OF behaviour)†	Association P value (OF-gene expression)‡
102852_at	Cadherin 2	<i>Cdh2</i>	pi	-1.72	0.95	7.7×10^{-4}
161603_r_at	Erythrocyte protein band 4.1-like 4a	<i>Epb4.1l4a</i>	pi	-3.98	0.89	2.5×10^{-2}
93268_at	Glyoxalase 1	<i>Glo1</i>	am, ci, bn, hi, hy , pa	-2.32	0.97	2.6×10^{-5}
93269_at	Glyoxalase 1	<i>Glo1</i>	am , ci, bn, hi, hy, pa	-2.53	0.94	7.8×10^{-5}
101044_at	Delta-aminolevulinatase dehydratase	<i>Alad</i>	hi, pa	-2.17	0.84	6.0×10^{-5}
160646_at	Glutathione reductase 1	<i>Gsr</i>	am, ci	-2.83	0.85	2.6×10^{-3}
101371_at	Cleavage and poly-adenylation specific factor 4	<i>Cpsf4</i>	hi	-1.90	0.80	5.2×10^{-4}
97560_at	Prosaposin	<i>Psap</i>	pa	-1.73	0.80	2.4×10^{-4}
102808_at	Voltage-gated sodium channel type Iβ	<i>Scn1b</i>	pi	-2.02	0.77	1.5×10^{-3}
101929_at	Dynein light chain 2	<i>Dlc2</i>	pa	-1.85	0.76	3.2×10^{-2}
92539_at	S100 calcium binding protein A10	<i>S100a10</i>	hy	1.80	-0.76	2.0×10^{-3}
101289_f_at	Kallikrein 21	<i>Klk21</i>	pi	6.74	-0.77	3.0×10^{-2}
101587_at	Epoxide hydrolase 1	<i>Ephx1</i>	hy	2.74	-0.78	5.6×10^{-3}
92783_at	Growth hormone	<i>Gh</i>	pa	5.20	-0.80	2.9×10^{-1}
103918_at	Solute carrier family 15 member 2	<i>Slc15a2</i>	ci	4.27	-0.80	2.6×10^{-6}
92546_r_at	Prostaglandin D2 synthase	<i>Ptgds</i>	bn , pa	2.67	-0.82	3.1×10^{-4}
100719_f_at	Kallikrein 16	<i>Klk16</i>	pi	5.54	-0.83	1.3×10^{-3}
98525_f_at	Erythroid differentiation regulator 1	<i>Erd1</i>	hi, hy	2.74	-0.87	3.5×10^{-2}
96215_f_at	cDNA clone MGC:67258	<i>Erd1</i>	hi	3.75	-0.98	3.1×10^{-3}

* Average fold change for the C57BL/6J and FVB/NJ versus A/J and DBA/2J comparisons. Value shown is the average over all tissues showing differential expression.

† In the case of multiple tissues, the most significant value is shown (for the tissue in bold).

‡ Based on the linear mixed-effects model analysis. am, amygdala; bn, bed nucleus of the stria terminalis; ci, cingulate cortex; hi, hippocampus; hy, hypothalamus; pa, periaqueductal grey; pi, pituitary; OF, open-field test.

of the cingulate cortex of C57BL/6J and 129S6/SvEvTac mice to overexpress the corresponding genes *in vivo*. These strains were selected because they are widely used in neurobiological research, with C57BL/6J representing a non-anxious strain and 129S6/SvEvTac representing an anxious strain. Injected animals were tested in the open-field test (Fig. 2d–e and data not shown). After testing, mice were allowed to recover for a week, killed, and their brains removed for immunohistochemical and *in situ* hybridization analysis. We confirmed transgene expression associated with stereotaxic injection by *in situ* hybridization (Supplementary Fig. 3b–c).

Overexpression of *Glo1* in the cingulate cortex of the anxious 129S6/SvEvTac strain further enhanced the anxiety-related phenotype. The *Glo1*-expressing mice spent 12% more time near the walls in the open-field chamber compared to the GFP-expressing controls ($P = 0.016$; Fig. 2d). This effect was evident as early as five weeks after injection. Similarly, 129S6/SvEvTac mice overexpressing *Gsr* in the cingulate cortex were more anxious than GFP-expressing controls, although the effect was on the border of statistical significance ($P = 0.054$; Fig. 2d). The less-anxious C57BL/6J mice injected with the *Gsr* lentivirus also showed an increase in anxious behaviour, spending 16% more time near the walls in the open-field chamber compared to GFP-expressing controls ($P = 0.003$; Fig. 2e). However, overexpression of *Glo1* in the C57BL/6J background did not increase the anxiety-related behaviour compared to GFP controls ($P = 0.212$; Fig. 2e). The behaviours of the three groups (*Glo1*-, *Gsr*- and GFP-expressing animals) were significantly different at five weeks after injection in 129S6/SvEvTac mice ($P = 0.047$), and at seven weeks after injection in C57BL/6J mice ($P = 0.040$), as shown by a Kruskal–Wallis non-parametric analysis of variance (ANOVA).

To further prove that the expression level of these genes modulates anxious behaviour, we tested whether inhibition of *Glo1* gene expression led to a decrease in anxiety-like behaviour using lentiviral vectors that expressed an siRNA (small interfering RNA) against *Glo1* (siGlo1). A control vector was used that expressed an siRNA against the human *p53* gene (sihp53)²³, which has been shown not to affect the expression of mouse *p53* (Supplementary Fig. 4; O.S. and I.M.V., unpublished results). The 129S6/SvEvTac and C57BL/6J strains of mice were injected with either a virus expressing siGlo1 or sihp53. Five weeks later, animals were tested using the open-field test. The 129S6/SvEvTac mice injected with siGlo1 virus spent 49% more time in the middle of the chamber compared with control animals injected with the sihp53 virus ($P = 0.036$; Fig. 2f). Likewise, C57BL/6J mice

injected with siGlo1 virus spent 38% more time in the middle of the chamber compared with control animals injected with the sihp53 virus ($P = 0.0002$; Fig. 2f), indicating that inhibition of *Glo1* expression in the cingulate cortex reduces levels of anxiety-like behaviour. We confirmed transgene expression associated with stereotaxic injection by visualizing GFP expression associated with lentiviral infection (Supplementary Fig. 3d).

The results of our lentivirus experiments show that overexpression of either *Glo1* or *Gsr* in the cingulate cortex increases, while inhibition of *Glo1* expression by siRNA decreases, the level of anxiety-like behaviour of mice. These results strongly support the hypothesis that changes in the expression levels of *Glo1* and *Gsr* in the brain lead to a significant effect on anxiety-related behaviour, and establish a causal role for these genes, which are both part of a pathway that regulates oxidative stress, in the genesis of anxiety-like behaviour.

We have shown that gene expression profiles of specific brain regions of anxious and non-anxious mice differ significantly. Our expression-based approach is expected to complement traditional QTL (quantitative trait loci) mapping: genes with expression levels that are correlated with the trait of interest and physically reside in close proximity to a QTL for the trait are good candidates for genes directly responsible for the QTL^{24,25}. In fact, several of our candidate genes reside within chromosomal regions with identified QTLs for anxiety-related behaviour^{26,27} (Supplementary Table 2). The newly identified genes should further our understanding of the specific genes, pathways and mechanisms that are important for the regulation of normal and pathological anxiety in mice and humans.

METHODS

Animals. Seven-week-old male mice were obtained from the Jackson Laboratory (A/J, BALB/cByJ, C3H/HeJ, C57BL/6J, DBA/2J, FVB/NJ and B6AF1/J) or from Taconic Farms (129S6/SvEvTac). AB6F₁ animals were bred at the Salk Institute using parental animals derived from the Jackson Laboratory. Animals were singly housed for one week before behavioural testing or dissections. All animal procedures were approved by the Salk Institute for Biological Studies institutional animal care and use committee. Different animals were used for behavioural testing and gene expression profiling in order to measure baseline gene expression differences.

Behavioural testing. Anxiety-related behaviour was measured using the light–dark box test and the open-field test (see the Supplementary Methods for details).

Tissue collection and RNA preparation. Animals were killed by cervical

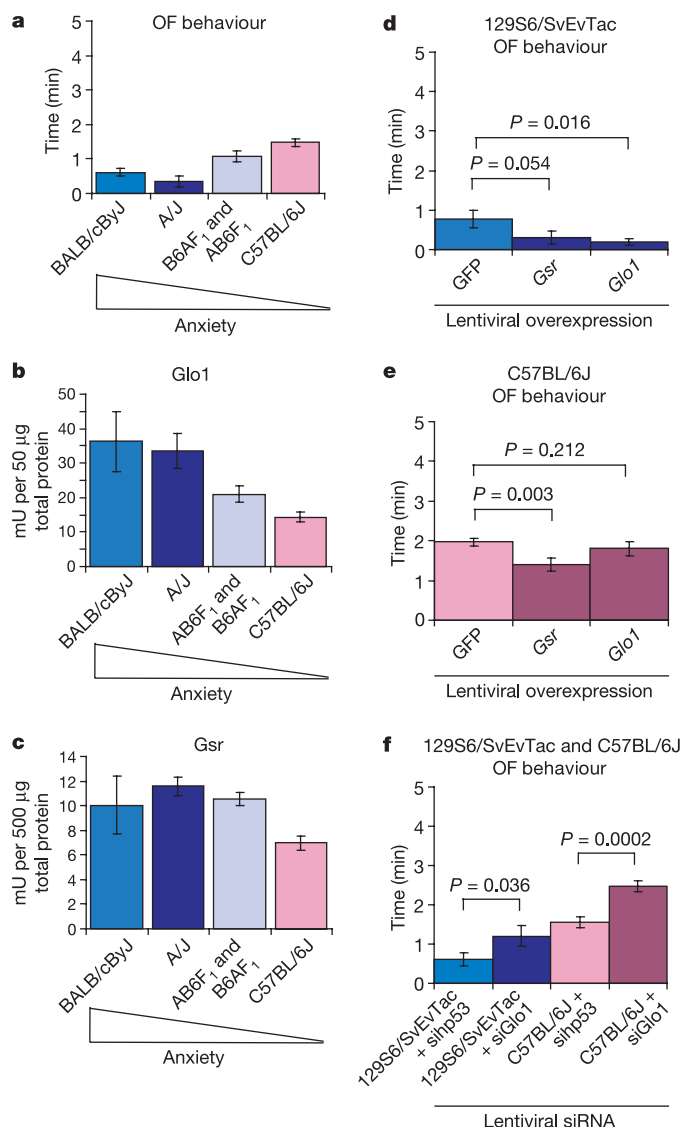


Figure 2 | Glyoxalase 1 (Glo1) and glutathione reductase 1 (Gsr) regulate anxiety-like behaviour in inbred mouse strains. **a**, Open-field (OF) behaviour. Mouse strains are shown on the x axis. Time spent in the middle of the open-field chamber is shown on the y axis. Values are mean $\pm$ s.e.m. **b**, Glo1 and **c**, Gsr brain enzyme activity (mean of two to four animals $\pm$ s.d.). See the Supplementary Methods for a description of the units. **d–f**, Open-field behaviour of *Glo1*-, *Gsr*- or GFP-overexpressing 129S6/SvEvTac mice five weeks after injection of the lentivirus (**d**); *Glo1*-, *Gsr*- or GFP-overexpressing C57BL/6J mice seven weeks after injection (**e**); and siGlo1- or sihp53-expressing 129S6/SvEvTac and C57BL/6J mice five weeks after injection (**f**). In each case, the x axis shows the name of the injected lentivirus. Time spent in the middle of the open-field chamber is shown on the y axis. Values are mean $\pm$ s.e.m. *P* values calculated using a one-tailed Student's *t*-test.

dislocation. All dissections were performed between 11.00–17.00 h on a Petri dish filled with ice using a dissection microscope. The dissected brain regions for gene expression analysis included the amygdala, cingulate cortex, hypothalamus, hippocampus, pituitary, periaqueductal grey and bed nucleus of the stria terminalis. Hippocampus samples were directly frozen on dry ice and stored at -80°C . The smaller brain structures were collected in RNA Later buffer (Ambion), and samples from 2–5 animals were pooled and stored at -80°C . The extraction of total RNA from the tissues was performed using the TRIzol reagent (Invitrogen) according to the manufacturer's instructions. Only samples with an absorbance ratio at 260 nm/280 nm (A_{260}/A_{280}) greater than 2.0 in TE buffer were used for further experiments.

Microarray experiments. Gene expression levels were measured using the

Murine Genome U74Av2 arrays (Affymetrix). Bed nucleus of the stria terminalis, hippocampus, hypothalamus, periaqueductal grey and pituitary gland samples were labelled using 10 µg of total RNA as the starting material. Owing to the small size of amygdala and cingulate cortex, samples from these tissues were labelled using 50 ng of total RNA as the starting material, using two rounds of complementary DNA synthesis and *in vitro* transcription (IVT). Labelling of samples, hybridization and scanning were performed as described²⁸. Two-round labelling was performed using the MessageAmp kit (Ambion) according to the manufacturer's instructions, with the exception that the second IVT was done using the Enzo BioArray high yield RNA transcript labelling kit (Enzo Life Sciences).

Data analysis. See the Supplementary Methods for further details concerning the analysis of differentially expressed genes and the determination of reproducibility between measurements, as well as the regression analysis between the behavioural results and enzyme activity levels.

Quantitative RT-PCR. PCR reactions were done using the SYBR Green master mix (Applied Biosystems) in an ABI Prism SDS 7900 HT machine (Applied Biosystems) as described in the Supplementary Methods.

Enzyme activity assays. Eight-week-old mice were killed by decapitation and their cortex, hippocampus and striatum dissected under a dissection microscope, frozen on dry ice, and stored at -80°C . The enzyme activity levels of Alad, Glo1 and Gsr were determined as described in the Supplementary Methods.

Lentivirus-mediated gene transfer. Plasmids were constructed for the production of lentiviral vectors that expressed either *Glo1* or *Gsr* with a carboxy-terminal HA-tag, or GFP, in the overexpression experiment. We sequenced the cDNA of *Glo1* and *Gsr* in order to find single nucleotide polymorphisms between the strains (see the Supplementary Methods and Supplementary Information). For the overexpression experiment, a variant of *Glo1* from the A/J strain was cloned. For the siRNA experiment, lentiviral vectors were constructed that expressed siRNA against *Glo1* (siGlo1) or human *p53* (sihp53) from the human H1-RNA promoter as described before (O.S. and I.M.V., unpublished results and ref. 23) (Supplementary Fig. 3a). Further details about virus production are given in the Supplementary Methods. A total of 50 129S6/SvEvTac and 50 C57BL/6J male mice were obtained from Taconic Farms or the Jackson Laboratory, respectively, at five weeks of age, and housed five mice per cage. After one week of acclimatization, mice were injected bilaterally with 1 µl (1.1×10^6 transducing units) of either HA-*Glo1*, HA-*Gsr*, GFP, siGlo1 or sihp53 virus (ten animals of both strains per construct) into the cingulate cortex using a stereotaxic frame. The stereotaxic coordinates were: 1.4 mm rostral to bregma, 0.5 mm lateral to midline, and 1.5 mm ventral from the dural surface. Four weeks after injection, mice were separated into individual cages. A few animals died after the injections, and the final number of animals used for further experiments are detailed in the Supplementary Methods. The open-field behavioural test was conducted five weeks and seven weeks after injection in the case of the overexpression experiment, and five weeks after injection in the case of the siRNA experiment. Mice were allowed to recover for a week, after which time they were killed and their brains were collected for the immunohistochemical or *in situ* hybridization analysis (see the Supplementary Methods for details). **Software tools.** Further details on the TeraGenomics microarray analysis tool are available at <http://www.teragenomics.com>. The Bullfrog software can be downloaded from <http://www.barlow-lockhartbrainmapnimhgrant.org/>.

Received 18 April; accepted 23 September 2005.

Published online 23 October 2005.

- Crawley, J. N. *et al.* Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. *Psychopharmacology (Berlin)* **132**, 107–124 (1997).
- Carter, T. A. *et al.* Chipping away at complex behaviour: transcriptome/phenotype correlations in the mouse brain. *Physiol. Behav.* **73**, 849–857 (2001).
- Fernandes, C. *et al.* Hippocampal gene expression profiling across eight mouse inbred strains: towards understanding the molecular basis for behaviour. *Eur. J. Neurosci.* **19**, 2576–2582 (2004).
- Belzung, C. & Griebel, G. Measuring normal and pathological anxiety-like behaviour in mice: a review. *Behav. Brain Res.* **125**, 141–149 (2001).
- Rogers, D. C. *et al.* Use of SHIRPA and discriminant analysis to characterise marked differences in the behavioural phenotype of six inbred mouse strains. *Behav. Brain Res.* **105**, 207–217 (1999).
- Bouwknegt, J. A. & Paylor, R. Behavioral and physiological mouse assays for anxiety: a survey in nine mouse strains. *Behav. Brain Res.* **136**, 489–501 (2002).
- Panksepp, J. Towards a general psychobiological theory of emotions. *Behav. Brain Sci.* **5**, 407–468 (1982).
- Le Douarin, J. Fear and the brain: where have we been, and where are we going? *Biol. Psychiatry* **44**, 1229–1238 (1998).

9. Camargo, E. E. Brain SPECT in neurology and psychiatry. *J. Nucl. Med.* **42**, 611–623 (2001).
10. Gray, J. A. & McNaughton, N. *The Neuropsychology of Anxiety* (Oxford Univ. Press, Padstow, 2000).
11. Hovatta, I. & Barlow, C. in *Databasing the Brain: From Data to Knowledge* (eds Koslow, S. H. & Subramaniam, S.) 169–190 (Wiley, 2005).
12. Kuloglu, M. *et al.* Antioxidant enzyme activities and malondialdehyde levels in patients with obsessive-compulsive disorder. *Neuropsychobiology* **46**, 27–32 (2002).
13. Kuloglu, M. *et al.* Antioxidant enzyme and malondialdehyde levels in patients with panic disorder. *Neuropsychobiology* **46**, 186–189 (2002).
14. Bilici, M. *et al.* Antioxidative enzyme activities and lipid peroxidation in major depression: alterations by antidepressant treatments. *J. Affect. Disord.* **64**, 43–51 (2001).
15. Yao, J. K. *et al.* Oxidative damage and schizophrenia: an overview of the evidence and its therapeutic implications. *CNS Drugs* **15**, 287–310 (2001).
16. Thornalley, P. J. Glyoxalase I—structure, function and a critical role in the enzymatic defence against glycation. *Biochem. Soc. Trans.* **31**, 1343–1348 (2003).
17. Chen, F. *et al.* Role for glyoxalase I in Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **101**, 7687–7692 (2004).
18. Junaid, M. A. *et al.* Proteomic studies identified a single nucleotide polymorphism in glyoxalase I as autism susceptibility factor. *Am. J. Med. Genet.* **131A**, 11–17 (2004).
19. Tafti, M. *et al.* Deficiency in short-chain fatty acid β -oxidation affects theta oscillations during sleep. *Nature Genet.* **34**, 320–325 (2003).
20. Kromer, S. A. *et al.* Identification of glyoxalase-I as a protein marker in a mouse model of extremes in trait anxiety. *J. Neurosci.* **25**, 4375–4384 (2005).
21. Naldini, L., Blomer, U., Gage, F. H., Trono, D. & Verma, I. M. Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc. Natl Acad. Sci. USA* **93**, 11382–11388 (1996).
22. Miyoshi, H., Blomer, U., Takahashi, M., Gage, F. H. & Verma, I. M. Development of a self-inactivating lentivirus vector. *J. Virol.* **72**, 8150–8157 (1998).
23. Tiscornia, G., Singer, O., Ikawa, M. & Verma, I. M. A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. *Proc. Natl Acad. Sci. USA* **100**, 1844–1848 (2003).
24. Sandberg, R. *et al.* Regional and strain-specific gene expression mapping in the adult mouse brain. *Proc. Natl Acad. Sci. USA* **97**, 11038–11043 (2000).
25. Karp, C. L. *et al.* Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma. *Nature Immunol.* **1**, 221–226 (2000).
26. Gershenfeld, H. K. & Paul, S. M. Mapping quantitative trait loci for fear-like behaviors in mice. *Genomics* **46**, 1–8 (1997).
27. Turri, M. G., Datta, S. R., DeFries, J., Henderson, N. D. & Flint, J. QTL analysis identifies multiple behavioural dimensions in ethological tests of anxiety in laboratory mice. *Curr. Biol.* **11**, 725–734 (2001).
28. Zapala, M. A. *et al.* Adult mouse brain gene expression patterns bear an embryologic imprint. *Proc. Natl Acad. Sci. USA* **102**, 10357–10362 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Information Management Consultants and Teradata (NCR) for donating and programming of the TeraGenomics database; Dan Lockhart, M. Zapala and N. Schork for software development and statistical analysis; B. Stoven for brain dissections; J. Airas for plasmid cloning; N. Tonnu for lentivirus production; F. Bloom, J. Reilly, W. Young, W. Vale and T. Carter for their insight; and members of the Barlow laboratory for discussions and technical assistance. This work was supported by grants from NIMH (to C.B. and D.J.L.), NINDS (to C.B.) and the Academy of Finland (to I.H.).

Author Contributions D.J.L. and C.B. conceived of and initiated the project. I.H., D.J.L. and C.B. designed the research. I.H. and R.S.T. performed the microarray, enzyme activity, sequencing and real-time qPCR experiments. I.H. and R.H. performed the behavioural analyses and lentivirus injections. I.H., J.M.R., J.A.E. and C.B. designed and J.M.R. performed the *in situ* hybridization experiments. I.H., R.A.M., O.S., I.M.V. and C.B. designed the lentivirus experiment, and R.A.M., O.S. and I.M.V. contributed the lentivirus vectors. I.H., E.E.S., C.B. and D.J.L. analysed the data. I.H., E.E.S., D.J.L. and C.B. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) and are accessible through the GEO series accession number GSE3327. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.B. (cbarlow@braincellsinc.com).

Risk of severe asthma episodes predicted from fluctuation analysis of airway function

Urs Frey¹, Tanja Brodbeck¹, Arnab Majumdar², D. Robin Taylor³, G. Ian Town³, Michael Silverman⁴ & Béla Suki²

Asthma is an increasing health problem worldwide¹, but the long-term temporal pattern of clinical symptoms is not understood and predicting asthma episodes is not generally possible^{2,3}. We analyse the time series of peak expiratory flows, a standard measurement of airway function that has been assessed twice daily in a large asthmatic population during a long-term crossover clinical trial⁴. Here we introduce an approach to predict the risk of worsening airflow obstruction by calculating the conditional probability that, given the current airway condition, a severe obstruction will occur within 30 days. We find that, compared with a placebo, a regular long-acting bronchodilator (salmeterol) that is widely used to improve asthma control decreases the risk of airway obstruction. Unexpectedly, however, a regular short-acting β_2 -agonist bronchodilator (albuterol) increases this risk. Furthermore, we find that the time series of peak expiratory flows show long-range correlations that change significantly with disease severity, approaching a random process with increased variability in the most severe cases. Using a nonlinear stochastic model, we show that both the increased variability and the loss of correlations augment the risk of unstable airway function. The characterization of fluctuations in airway function provides a quantitative basis for objective risk prediction of asthma episodes and for evaluating the effectiveness of therapy.

Asthma is a chronic inflammatory disease of the airways. The complex interactions between endogenous and environmental factors result in a highly variable pattern of airway obstruction over time^{5–7}. Fluctuations in airway calibre result in episodic symptoms of wheeze, dyspnoea or cough⁸. The interpretation and prediction of such fluctuations are difficult not only because environmental stimuli are not always recognizable and are difficult to quantify, but also because the correlation between stimuli and symptoms is exceedingly poor^{2,9}. Short-term variability of airway resistance and its distribution are sensitive to the asthmatic condition⁷. Predicting the risk of asthma episodes requires a detailed knowledge of the temporal patterns of the fluctuations in airway function; however, the long-term temporal properties of airway obstruction in asthma has not been studied. Such predictions should be valuable for individuals, as well as for evaluating therapeutic interventions in clinical trials.

Here we investigate the temporal pattern and their predictive utility of airway function characterized by peak expiratory flows (PEFs), defined as the maximum flow measured during a single forced expiratory manoeuvre, in 80 asthmatic subjects (Methods). In particular, we examine whether the statistical and correlation properties of the time series of PEF recordings can be used to predict the risk of subsequent exaggeration of airway instability. Using data from a previously published, randomised, placebo-controlled, double-blind, crossover study⁴, we analyse serial, twice-daily PEFs, together

with daily asthma symptom scores, obtained in three 6-month treatment periods. The trial compared the effects of a regular short-acting β -agonist treatment (albuterol, 400 μ g four times daily; 'SA period'), and a regular long-acting β -agonist treatment (salmeterol, 50 μ g twice daily; 'LA period') with those of a matching placebo (PL period). Subjects had similar mean daily doses of inhaled corticosteroids in all treatment periods (median ranged between 500 and 548 μ g d⁻¹ of beclomethasone or equivalent).

The composite daily clinical asthma scores⁴ averaged and normalized per day (Asc) were significantly lower (paired *t*-test, $P < 0.05$) in the LA period, implying better asthma control than in the SA and PL periods. Similarly, individuals had significantly more wheezing episodes (N_{wheeze}) in the SA and PL periods than in the LA period ($P = 0.05$). Furthermore, the subjects required significantly more additional 'on-demand' reliever (albuterol) in the SA and PL treatment periods than in the LA period ($P < 0.05$). In the SA period, the number of severe asthma episodes in a given 6-month treatment period (N_{episodes}), defined by the need for oral corticosteroid treatment, was more than twice (13) that in the LA (6) and PL (5) periods. These data indicate markedly less stable airway function in the SA period.

The individual time series of the 300 consecutive twice daily PEF values show substantial fluctuations especially in the SA period (Fig. 1a). As compared with PL, the distribution of PEF in the LA period shifts to higher values consistent with improved airway function (Fig. 1b). The distribution of PEF in the SA period becomes skewed, however, reaching significantly lower PEF values than in the PL period, a sign of increased instability of airway function. All individuals show a similar pattern (Table 1). The means of the individual PEF series averaged within each group are significantly higher (paired *t*-test and signed ranks test, $P < 0.001$) in the LA than in either the SA or PL period, whereas those in the SA are higher than in the PL period. Conversely, the variability (coefficient of variation; CV), of the PEF series is significantly higher in the SA than in either the LA or PL period (paired *t*-tests, $P \leq 0.005$), whereas the CV in the LA is lower than in the PL period ($P = 0.001$). These findings are in accord with previous observations of diurnal effects of short-acting β_2 -agonist drugs^{4,10} and are probably a consequence of the long night time interval without drugs together with rebound bronchoconstriction. Mean skewness is higher in the LA than in the SA and PL periods (signed ranks test, $P < 0.05$).

To determine whether the variability of PEF is accompanied by long-range correlations, we calculate the detrended fluctuation function $F(n)$ (ref. 11) from the PEF series (Methods). We find that $F(n)$ follows a power law functional form, $F(n) \approx n^\alpha$ (where n is the window size), as judged by the linear increase of $F(n)$ on a double logarithmic graph in all three treatment periods (Fig. 1c). The exponent α characterizes the correlation properties of the PEF series:

¹Paediatric Respiratory Medicine, Department of Paediatrics, University Hospital of Berne, Inselspital 3010 Berne, Switzerland. ²Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, USA. ³Faculty of Medicine, Dunedin/Christchurch Schools of Medicine, University of Otago, New Zealand. ⁴Institute for Lung Health, Leicester University, Leicester LE2 7LX, UK.

for $\alpha = 0.5$ the series is uncorrelated, whereas for increasingly higher values of α it shows increasingly stronger long-range correlations. Both treatments preserve the long-range correlations, but in this individual (Fig. 1c), α decreases by 17% in the SA period and increases by 15% in the LA period. Thus, SA not only fails to increase the mean PEF (Fig. 1a), it increases the variability of PEF (Fig. 1b) and alters the correlations of PEF to become more random such that past PEF values have less effect on the current or future values. Conversely, in the LA period there is a stronger correlation between past and future values of PEF, strengthening the memory of the system. We find evidence of long-range correlations in all individuals (Table 1), with trends in the LA and SA periods similar to those in Fig. 1c.

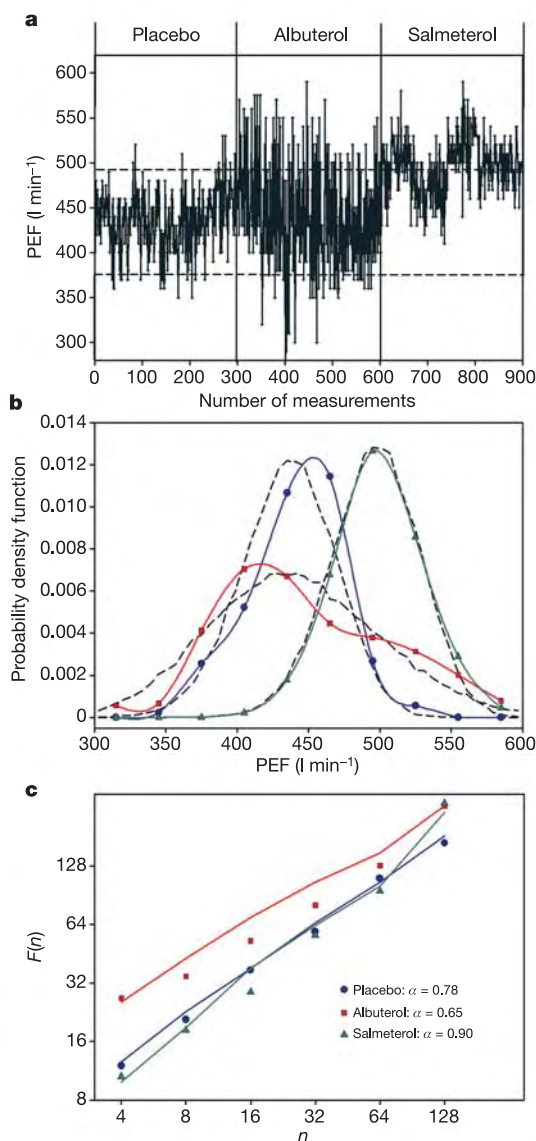


Figure 1 | Representative example of PEF time series in an asthmatic individual. **a**, Examples of time series of 300 PEF measurements in each of three treatment periods with placebo (PL, blue), short-acting albuterol (SA, red) and long-acting salmeterol (LA, green) in an asthmatic individual. The upper and lower dashed lines indicate 80% and 60%, respectively, of the predicted normal values. **b**, Distributions of PEFs corresponding to the data in **a** (unbroken lines) and obtained from simulations (broken lines) with a nonlinear stochastic model (Methods). In this case, the mean PEF values for PL, SA and LA are 436, 441 and 496 l min⁻¹ with s.d. values of 33, 61 and 31 l min⁻¹ and skewness values of -0.19, 0.24 and -0.17, respectively. **c**, Detrended fluctuation functions, $F(n)$, of the time series in **a** (symbols) and from the nonlinear model¹¹ (unbroken lines; Methods).

Table 1 | PEF in the three treatment periods

PEF	Mean (l min ⁻¹)	Median (l min ⁻¹)	CV (%)	Skewness	α
Salmeterol (LA)	447 (99)	448 (100)	5.6 (2.3)	-0.51 (1.08)	0.82 (0.15)
Albuterol (SA)	429 (98)	428 (99)	7.7 (3.2)	-0.18 (0.76)	0.73 (0.16)
Placebo (PL)	417 (99)	419 (100)	6.8 (2.6)	-0.26 (0.74)	0.78 (0.19)

Means, medians, CVs, skewness values and correlation exponents (α) were calculated from PEF time series and averaged for the three treatment periods (PL, SA, LA) of asthmatic subjects. Numbers in parentheses are s.d. values for the group. The skewness shows a range of values in each group; however, it is a less important determinant of severe airway obstruction than is CV or α (see Fig. 4).

Regarding the mean values, although α tended to be higher in the LA than in the PL period, it decreased significantly in the SA as compared with the PL period ($P < 0.001$). In the PL period, α is directly correlated with baseline lung function (linear regression: $\alpha = 0.0032\text{PEF}_{\text{pred}} + 0.532$, where PEF_{pred} is the percentage predicted PEF value; $P < 0.02$) and is inversely related to N_{wheeze} in that period ($\alpha = -0.00037N_{\text{wheeze}} + 0.819$; $P < 0.05$), but not to asthma symptom scores. Similar trends are seen in the SA and LA periods. Thus, lower values of α generally reflect more severe airflow obstructions. However, the correlation between α and airway function is more complex: the closer α is to the median value of 0.78, the smaller is the change in α ($\Delta\alpha$) with treatment (Fig. 2). Of greater clinical relevance is the fact that $\Delta\alpha$ correlates with the improvement in the clinical asthma score ΔAsc (multiple linear regression adjusting for treatment, SA or LA: coefficient, 0.305; 95% confidence interval, 0.171–0.439; $P < 0.001$) but not with ΔN_{wheeze} . Whereas α and CV are not significantly correlated, $\Delta\alpha$ and ΔCV are highly correlated in the PL period (coefficients: 0.026, 0.014–0.038; $P < 0.001$).

These results imply that both α and the CV should be related to the risk of acute asthma episodes with significant airflow obstruction. To assess this risk, we calculate the conditional probability (π) that a significant deterioration in airway obstruction, defined as $\text{PEF} < 80\%$ (moderate) or $\text{PEF} < 60\%$ (severe)¹² of the age- and height-predicted normal values, occurs within a certain time period given the current value of PEF (Methods). The probability that a moderate airflow obstruction occurs within one month decreases from nearly 100% at low initial values of PEF ($< 200 \text{ l min}^{-1}$) to between 10 and 30% for high initial values of PEF ($> 550 \text{ l min}^{-1}$) depending on the treatment period (Fig. 3). The probability that an episode of severe airway obstruction occurs within the same time frame is similar but with lower values. We also calculate the

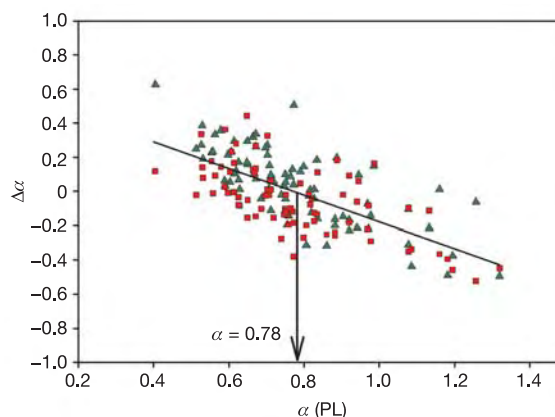


Figure 2 | Change in α as a function of α in the placebo treatment period. Shown is the change in α ($\Delta\alpha$) from the placebo treatment period (PL) to the salmeterol (LA, green) or albuterol (SA, red) treatment period as a function of the value of α in the PL period. Unbroken line represents the regression line for all data ($\Delta\alpha = -0.79\alpha + 0.61$; $P < 0.001$). Minimal change in α occurs when α in the PL period is 0.78.

relationship between π and PEF on an individual basis. For any value of PEF, regular long-acting salmeterol significantly decreases the risk as compared with both placebo (paired *t*-test, $P < 0.004$) and regular short-acting albuterol ($P < 0.02$). Unexpectedly, however, albuterol increases the risk of future moderate or severe airflow obstruction beyond that seen with placebo especially for near normal values of PEF (Fig. 3b). Loss of asthma control observed with regular short-acting β -adrenergic agonist treatment has been described^{13–15}. The predominant view is that this loss is due to β -adrenergic desensitization¹³, but our results suggest alternatively that the short-acting agonist treatment leads to significantly increased variability and loss of predictability of airway function.

To assess the separate effects of the distribution and correlation properties of the PEF series on the risk π , we introduce a nonlinear stochastic model of the PEF fluctuations (Methods). The input to the model is gaussian white noise, and the output of the model is a time series that matches both the distribution (Fig. 1b) and the correlation properties (Fig. 1c) of PEF. We then generate several time series, each containing 50,000 points, for various combinations of model parameters and calculate the corresponding π . We find that π decreases with increasing α and is highly sensitive to the variance but less sensitive to the skewness of the model output (Fig. 4). We also average the individual probabilities π corresponding to a set of intervals of α in the PL group, which compare well with the model predicted range. These results indicate that the increased risk in the SA period is a consequence of the decreased α and increased CV of the PEF time series. Conversely, the decreased risk in the LA period

results primarily from the increased median PEF, the decreased CV, and perhaps also from an increase in α (Table 1, Fig. 3). Although CV seems to have a dominant effect on π (Fig. 4), the correlation properties of PEF are also important because, in the absence of correlations ($\alpha = 0.5$), π would be identical to regular probability and the curves in Fig. 3 would be independent of PEF. When PEF is high, therefore, a high value of α is beneficial because it lowers the risk. If PEF is low, however, a high α value is not beneficial because the PEF would tend to remain low. Achieving or maintaining stable asthma with a given treatment thus requires a high average PEF accompanied by strong correlations or a high value of α .

Our approach of estimating the conditional probability of moderate or severe airway obstruction and hence the risk of asthma exacerbations in individuals can have practical benefits for patient management. Currently, short-acting bronchodilators are the first line drug for as-required treatment of asthma symptoms. Our study suggests, however, that during regular use (four times daily with a long night time drug-free interval), short-acting bronchodilators do not decrease but even increase the risk of asthma episodes owing to the lack of maintained β -agonist effectiveness. By contrast, regular long-acting bronchodilators are more effective at stabilizing airway function over extended periods.

A possible interpretation of these results is that the PEF time series of normal subjects and individuals with mild and stable asthma show long-range baseline correlations. The physiological origins of these correlations are currently unknown. With increasing asthma severity, airways become hypersensitive to even apparently insignificant environmental factors such as small amounts of pollutants, allergens or minor viral infections, which may trigger serious and unexpected events such as an acute asthma attack caused by a sudden catastrophic collapse of large airway clusters¹⁶. Because the normal fluctuations in airway function are interrupted by these irregular and unrelated events, the baseline correlations are lost and the variability of PEF is increased.

Our results facilitate an understanding of unstable and stable

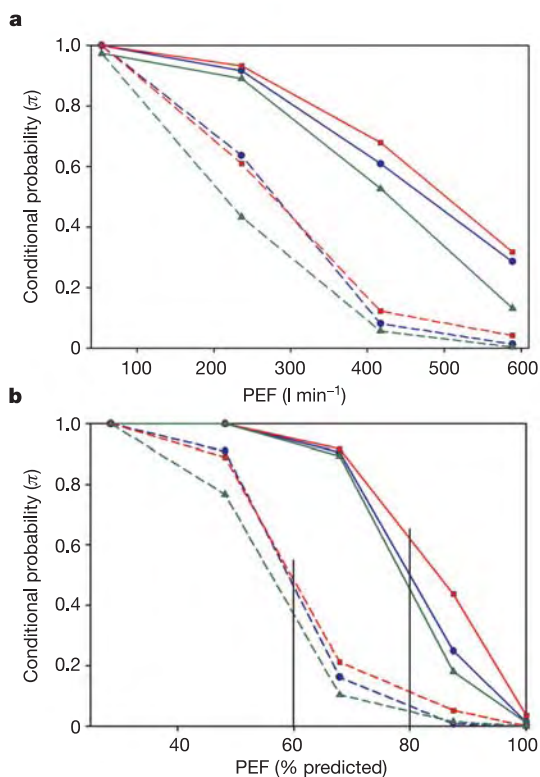


Figure 3 | Averaged probability that airway obstruction will occur as a function of PEF. **a**, Conditional probabilities (π) that an episode of airway obstruction will occur within 1 month as a function of the current value of PEF in the three treatment periods (PL, blue; SA, red; LA, green). Episodes of moderate or severe airway obstruction are defined as PEF $< 80\%$ (unbroken lines) and PEF $< 60\%$ (broken lines), respectively, of the age-predicted normal values of PEF. **b**, Conditional probabilities as a function of PEF presented as a percentage of age-predicted normative values. The risk π is the lowest in LA but, unexpectedly, for large values of the relative PEF, it is higher in SA than in PL.

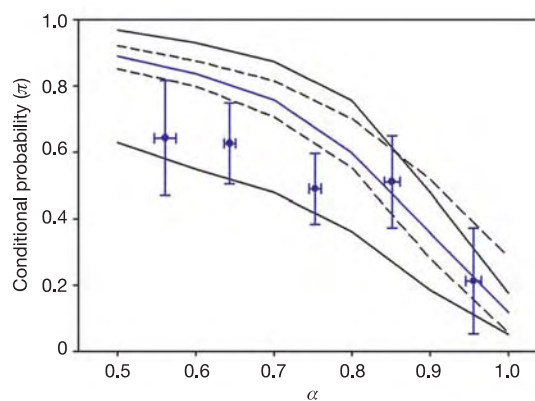


Figure 4 | Model predictions of the sensitivity of conditional probability. Shown are predictions of the sensitivity of the conditional probability (π) to determine that an episode of moderate airway obstruction will occur within 1 month when the current PEF is near its mean value. The mean (417 l min^{-1}), CV (6.8%) and skewness (-0.26) values are selected from Table 1 to match those of the placebo data (blue line). In each case, the mean and either the CV or the skewness is kept constant, and π is calculated as a function of α . For all parameter combinations, π decreases considerably with increasing α . The value of π is very sensitive to CV: the lower and upper unbroken black curves correspond to CV values of 5.8% and 7.8%, respectively. The skewness has a smaller effect: π decreases or increases only slightly when the skewness is halved or doubled, respectively (broken lines). For the experimental relationship between π and α (blue circles), we include data only for initial PEF values close to the mean PEF and present the mean $\pm$ s.e.m. (bars) of π corresponding to different ranges of α (from 0.5 to 0.6, 0.6 to 0.7, and so on).

asthma that can lead to a rational optimization of treatment. In addition, the quantitative approach proposed here can be applied to uncover dynamic patterns in the fluctuations of clinical symptoms of other complex chronic diseases of which asthma is a representative model.

METHODS

Subjects. The analysis was performed in a PEF time series of 80 non-smoking individuals (mean age, 42.4 yr; s.d., 11.8 yr; range, 19–64 yr) with persistent asthma. Baseline lung function and clinical symptom scores were assessed before the study. The three treatment periods of 6 months each were randomly sequenced. PEF, symptoms and 'as-required' bronchodilator use were recorded twice daily. Of the original 165 participants, 80 individuals with less than 3% of values missing are included in the analysis. This criterion is necessary to ensure that the correlation analysis can be carried out appropriately.

Correlation analysis and conditional probability. Correlation analysis was done with a detrended fluctuation function $F(n)$ (ref. 11). To calculate $F(n)$, the PEF series is integrated and divided into equally sized non-overlapping windows of length n . A linear regression line is fitted through the data in each window and the time series are locally detrended by subtracting the regression line from the data. The $F(n)$ is computed for each window of length n as the root-mean-square of the detrended and integrated PEF time series. The calculation is repeated for different values of n , and α is obtained as the slope of a straight line fit to $F(n)$ on a log–log plot.

The conditional probability (π) is calculated by first establishing five bins of the PEF values and a window of 60 points corresponding to 1 month. As the window moves along the time series in steps of one data point, the first point of the window determines the bin number k and a counter $N_1(k)$ is incremented. In addition, a separate counter, $N_c(k)$, is also incremented if any value of PEF (other than the first) in the window falls below the selected 80% or 60% threshold. The N_1 and N_c for all k are then summed for all individuals corresponding to a given condition and $\pi(k)$ is obtained as the ratio $N_c(k)/N_1(k)$.

Nonlinear stochastic model. We introduce a nonlinear stochastic block-structured model of the PEF fluctuations. We assume that the model is a cascade connection of a linear dynamic system (L) followed by a second order nonlinear system with no memory (N). The input to L is a zero mean gaussian white noise $x(t)$ with unit variance. To account for the power law correlations in the experimental data, the memory of L is modelled by a transfer function $H(f) \approx f^{\beta/2}$ where f is the frequency and the exponent β is related to the correlation exponent by $\alpha = (\beta + 1)/2$. The output $u(t)$ of L is a long-range correlated signal with exponent α and is led through N to produce a final output $y(t) = a_0 + a_1 u(t) + a_2 u(t)^2$. To match the mean (μ), variance (σ^2) and skewness (γ) of the experimental data, we solve for the higher order moments of y and approximate the coefficients as $a_2 = \sigma\gamma(1 + 0.0185\gamma^2)/6$, $a_1^2 = \sigma^2 - 2a_2^2$ and $a_0 = \mu - a_2$. Examples of the distributions and correlations of $y(t)$ are shown in Fig. 1b and c, respectively. Because the experimental γ is small (Table 1), $a_2 < a_1/20$, and hence the contribution of nonlinearity to the correlations of $y(t)$ is negligible, which we verified numerically. To obtain π from the model, we generate a time series of $y(t)$ containing 50,000 points and calculate π as described above. The number of data points needed for detrended fluctuation analysis is crucial because finite size effects can occur with short time series. Finite size error introduced by analysing 300 data points is less than 2–3% (ref. 17). Our simulated PEF series using 50,000 values provides results consistent with those extracted from our clinical records.

The model inputs and parameters could be interpreted as follows. The input

to L may represent the fluctuations in environmental stimuli (such as pollutants, allergens or viral infections) and internal stimuli (such as acute inflammation, exercise or hyperventilation). The correlation exponent β and the parameters a_0 , a_1 and a_2 may be related to airway closure, persistent inflammatory processes, increased airway smooth muscle sensitivity and airway wall remodelling⁸.

Received 25 June; accepted 31 August 2005.

1. Masoli, M., Fabian, D., Holt, S. & Beasley, R. The global burden of asthma: executive summary of the GINA Dissemination Committee report. *Allergy* **59**, 469–478 (2004).
2. Plaza, V., Serrano, J., Picado, C. & Sanchis, J. Frequency and clinical characteristics of rapid-onset fatal and near-fatal asthma. *Eur. Respir. J.* **19**, 846–852 (2002).
3. Zhang, J., Yu, C., Holgate, S. T. & Reiss, T. F. Variability and lack of predictive ability of asthma end-points in clinical trials. *Eur. Respir. J.* **20**, 1102–1109 (2002).
4. Taylor, D. R. et al. Asthma control during long-term treatment with regular inhaled salbutamol and salmeterol. *Thorax* **53**, 744–752 (1998).
5. Cookson, W. The alliance of genes and environment in asthma and allergy. *Nature* **402** (suppl.), B5–B11 (1999).
6. Holgate, S. T. Genetic and environmental interaction in allergy and asthma. *J. Allergy Clin. Immunol.* **104**, 1139–1146 (1999).
7. Que, C. L., Kenyon, C. M., Olivenstein, R., Macklem, P. T. & Maksym, G. N. Homeokinesis and short-term variability of human airway caliber. *J. Appl. Physiol.* **91**, 1131–1141 (2001).
8. Suki, B. & Frey, U. Temporal dynamics of recurrent airway symptoms and cellular random walk. *J. Appl. Physiol.* **95**, 2122–2127 (2003).
9. Olden, K. & Wilson, S. Environmental health and genomics: visions and implications. *Nature Rev. Genet.* **1**, 149–153 (2000).
10. Taylor, D. R. et al. Regular inhaled β agonist in asthma: effects on exacerbations and lung function. *Thorax* **48**, 134–138 (1993).
11. Peng, C. K. et al. Long-range anticorrelations and non-Gaussian behaviour of the heartbeat. *Phys. Rev. Lett.* **70**, 1343–1346 (1993).
12. Global strategy for asthma management and prevention (GINA), Publication No. 02-3659 7 (NHLBI, Bethesda, Maryland, USA, 2004).
13. Shore, S. A. & Drazen, J. M. β -Agonists and asthma: too much of a good thing? *J. Clin. Invest.* **112**, 495–497 (2003).
14. Anderson, H. R. et al. Bronchodilator treatment and death from asthma: case control study. *Br. Med. J.* **330**, 117–120 (2004).
15. O'Byrne, P. M. et al. Budesonide/formeterol combination therapy as both maintenance and reliever medication in asthma. *Am. J. Respir. Crit. Care Med.* **15**, 129–136 (2005).
16. Venegas, J. G. et al. Self-organized patchiness in asthma as a prelude to catastrophic shifts. *Nature* **434**, 777–782 (2005).
17. Cernelc, M., Suki, B., Reinmann, B., Hall, G. L. & Frey, U. Correlation properties of tidal volume and end-tidal O_2 and CO_2 concentrations in healthy infants. *J. Appl. Physiol.* **92**, 1817–1827 (2002).

Acknowledgements This work was supported by the Swiss NSF (U.F.) and the US NSF (B.S.).

Author Contributions U.F. and B.S. were the main investigators responsible for the study, including the analysis design; T.B. carried out large parts of the data analysis; A.M. calculated the stochastic model; D.R.T. and G.I.T. conducted the original clinical trial; M.S. contributed to study design and to writing the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to U.F. (urs.frey@insel.ch).

The microRNA *miR-196* acts upstream of *Hoxb8* and *Shh* in limb development

Eran Hornstein¹, Jennifer H. Mansfield¹, Soraya Yekta², Jimmy Kuang-Hsien Hu¹, Brian D. Harfe³, Michael T. McManus⁴, Scott Baskerville², David P. Bartel² & Clifford J. Tabin¹

MicroRNAs (miRNAs) are an abundant class of gene regulatory molecules (reviewed in refs 1, 2). Although computational work indicates that miRNAs repress more than a third of human genes³, their roles in vertebrate development are only now beginning to be determined. Here we show that *miR-196* acts upstream of *Hoxb8* and Sonic hedgehog (Shh) *in vivo* in the context of limb development, thereby identifying a previously observed but uncharacterized inhibitory activity that operates specifically in the hindlimb. Our data indicate that *miR-196* functions in a fail-safe mechanism to assure the fidelity of expression domains that are primarily regulated at the transcriptional level, supporting the idea that many vertebrate miRNAs may function as a secondary level of gene regulation.

Sonic hedgehog (Shh) is a key signal mediating anteroposterior polarity in both the fore- and hindlimb buds⁴. Retinoic acid (RA) signalling is required for Shh expression in the forelimb and the hindlimb^{5–8}. The transcription factor *Hoxb8* seems to mediate the induction of Shh by RA in the forelimb in that *Hoxb8* is upregulated as an immediate-early response to ectopic RA administered to the chick forelimb bud⁷, and ectopic *Hoxb8* expression in the anterior of the forelimb of a transgenic mouse leads to Shh expression⁹. Ectopic RA does not lead to *Hoxb8* induction in the hindlimb bud, however, owing to the presence of an unknown hindlimb-specific inhibitory activity¹⁰.

Reasoning that the unknown hindlimb inhibitory activity¹⁰ might be mediated by a small silencing RNA, we blocked miRNA processing by using a conditional knockout allele of *Dicer*, a key enzyme required for producing functional miRNAs from their precursors^{11,12}. *Dicer* activity can be specifically removed from the limb buds by using a conditional allele¹³ and a limb-specific *Prx1::cre* construct¹⁴ (Supplementary Fig. 1a), which recombine floxed alleles efficiently in the limb mesenchyme (Supplementary Fig. 1b). To test whether the inhibition of *Hoxb8* induction by RA in hindlimb buds is relieved by the removal of *Dicer* activity, hindlimbs from *Dicer*^{Δflox/Δflox} and wild-type mice at embryonic day 11.5 (E11.5) were cultured in the presence of RA. As in chick limbs, the presence of RA led to a marked upregulation of *Hoxb8* messenger RNA in the forelimb tissue of both wild-type and mutant animals (Fig. 1a, b), but not in wild-type hindlimbs (Fig. 1c). In *Dicer*^{Δflox/Δflox} hindlimbs, however, RA induced the expression of *Hoxb8* (Fig. 1d). As previously shown¹³, loss of *Dicer* activity does not affect the expression of other known patterning genes in the developing limb bud (Supplementary Fig. 1c). Thus, the previously uncharacterized inhibitory activity¹⁰ is lost in the absence of *Dicer*.

Dicer is crucial for the processing of hundreds of miRNAs and many siRNAs. To identify specific candidate miRNAs that could be

responsible for the hindlimb-specific inhibitory activity downstream of *Dicer*, we used microarray analysis¹⁵. Of the miRNAs that are expressed in the limb primordia, 12 were at least twofold more abundant in either the forelimb or the hindlimb bud (Fig. 2a and Supplementary Table 1). The most differentially expressed miRNA in the screen was *miR-196*, with an expression signal in the hindlimb exceeding by 20-fold that in the forelimb (Fig. 2a and Supplementary Table 1). Differential *miR-196* expression was verified in northern blot analyses of RNA isolated from forelimbs and hindlimbs of both chick and mouse (Fig. 2b) and was also consistent with the expression domain suggested by a transgenic reporter study¹⁶. Intriguingly, *Hoxb8* mRNA is a known target of *miR-196* *in vivo*^{16,17}. Therefore, we investigated whether *miR-196* might be the unknown hindlimb-specific activity preventing *Hoxb8* induction by RA.

First, to establish that *Hoxb8* is indeed an *in vivo* target of *miR-196* in the hindlimb, we carried out a modified 5' rapid amplification of

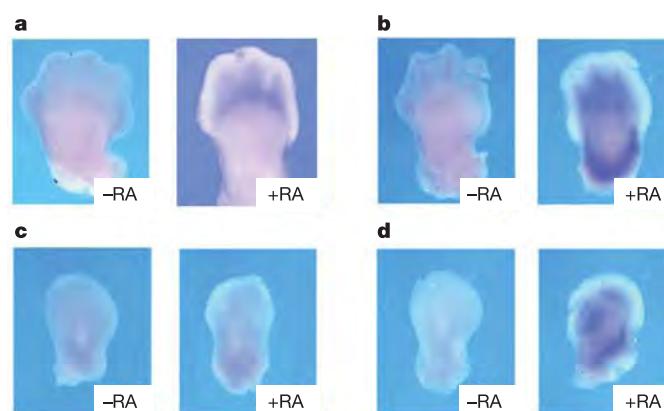


Figure 1 | Activity downstream of *Dicer* inhibits RA-induced expression of *Hoxb8* in mouse hindlimbs. **a**, E11.5 *Dicer*^{flox/+};*Prx1::Cre* (wild-type) forelimbs were cultured without RA (–RA), leading to no detection of *Hoxb8* (*n* = 6/6), or with 100 nM RA for 12 h (+RA), leading to induction of *Hoxb8* (*n* = 6/6). Expression of *Hoxb8* was detected by means of whole-mount *in situ* hybridization. **b**, E11.5 *Dicer*^{flox/flox}/*Dicer*^{flox/flox};*Prx1::Cre* (*Dicer* knockout) forelimbs subjected to the same treatment similarly resulted in induction of *Hoxb8* only in the presence of RA (*n* = 6/6 negative, without RA; 6/6 positive, with RA). **c**, Hindlimbs from the mice in **a** were cultured similarly, and RA failed to induce *Hoxb8* expression (*n* = 8/8 with RA; 8/8 without RA). **d**, Hindlimbs from the *Dicer* knockout mice in **b** were cultured similarly. Complete deletion of *Dicer* did not result in induction of *Hoxb8* in untreated hindlimbs (*n* = 5/5), but it enabled the accumulation of *Hoxb8* transcripts in RA-treated hindlimb mesenchyme (*n* = 6/6).

¹Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA. ²Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142, USA. ³Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, Gainesville, Florida 32610, USA. ⁴Department of Microbiology and Immunology, Diabetes Center, University of California at San Francisco, San Francisco, California 94143, USA.

complementary DNA ends (RACE) protocol, commonly used as an assay for miRNA-directed mRNA cleavage^{17,18}. By sequencing the 5' RACE products, we could determine whether any amplified *Hoxb8* degradation products were cleaved precisely at the predicted *miR-196*-binding site. We could easily observe *miR-196*-directed *Hoxb8* cleavage in the wild-type hindlimb, whereas *Hoxb8* cleavage in the forelimb tissue was barely seen (Fig. 3a, b). These data indicate that *Hoxb8* is indeed both transcribed, at a level detectable by polymerase chain reaction (PCR), and cleaved *in vivo* in the hindlimb.

In wild-type chick embryo, after 2.5 d of incubation *Hoxb8* is expressed in the neural tube and somites. *Hoxb8* is also expressed in the forelimb field, where it functions in inducing *Shh* during the early limb field stages (Fig. 3c). To test whether *miR-196* activity could attenuate *Hoxb8* expression at the early limb field (stage 16), we used a replication-competent viral expression system (RCAS). Our analysis showed that 26 h after *in ovo* injection of the virus RCAS::miR-196, *Hoxb8* expression was reduced throughout the embryo and, in particular, endogenous expression of *Hoxb8* in the forelimb field was markedly repressed (Fig. 3d).

We next addressed whether *miR-196* could be responsible for the inability of ectopic RA to induce *Hoxb8* in the hindlimb¹⁰. We implanted RA-soaked beads into wild-type chick forelimbs, which induced *Hoxb8* within 4 h (Fig. 3e). By contrast, parallel implantations failed (or were only marginally able) to induce *Hoxb8* in forelimb buds ectopically expressing *miR-196* (Fig. 3f). Misexpression of *miR-196* in the forelimb thus creates a situation that is reminiscent of wild-type hindlimb, in which endogenously high expression of *miR-196* leads to observable degradation of endogenous *Hoxb8* and correlates with an inability of RA to induce ectopic *Hoxb8*.

The *miR-196*-sensitivity of *Hoxb8* thus provides a compelling explanation for the inability of RA to induce *Hoxb8* in the hindlimb. In previous studies^{7,10}, RA and *Hoxb8* were placed upstream of *Shh* expression in the forelimb and, indeed, blocking endogenous RA activity resulted in a significant, albeit incomplete, downregulation of endogenous *Shh* expression^{7,10}. If the *miR-196*-sensitivity of *Hoxb8* expression were truly involved in mediating RA-induced expression of *Shh* in the forelimb bud, then *Shh* expression itself should be downregulated on the introduction of *miR-196* into the forelimb. Indeed, when chick embryos were analysed 2 d after viral misexpression of *miR-196* in the right limb field, endogenous *Shh* was consistently downregulated (Fig. 4a, compare with 4b). Other genes, not described to be downstream of *Hoxb8* in the limb mesenchyme,

were not affected by misexpression of *miR-196*, suggesting that this was a specific effect (Supplementary Fig. 2). To quantify the effect of *miR-196* on *Shh* levels, we infected chick embryos as above and assayed them 2 d later by quantitative real-time PCR. *Shh* expression was decreased in the *miR-196*-infected forelimb to roughly a third of the level seen in wild-type limbs (Fig. 4c).

We also checked whether ectopic misexpression of *miR-196* would block RA-induced ectopic expression of *Shh*. When RA-soaked beads were implanted into wild-type chick forelimb for 36 h, an anterior domain of ectopic *Shh* was induced⁴ (Fig. 4d); however, in *miR-196*-infected limbs, *Shh* expression was blocked or diminished and more diffuse (Fig. 4e). Although *Shh* was repressed by *miR-196* misexpression in the forelimb, the expression of *Shh* in the hindlimb was not affected by the same manipulation (Fig. 4f, g). This difference highlights the rather unexpected conclusion that independent pathways control *Shh* expression in the forelimb and the hindlimb (Fig. 4h), which may be explained by a dual role for Hox genes in specifying forelimb versus hindlimb identity and in regulating *Shh* expression. After *Hoxb8* and other related Hox genes evolved to specify forelimb-specific morphology, a different, *Hoxb8*-independent, mechanism of regulating *Shh* downstream of RA had to evolve for the hindlimb.

Despite the evidence presented here and elsewhere⁹ supporting a role for *Hoxb8* in regulating *Shh* in the forelimb, it has been reported that even the removal of all three Hox8 paralogues has no effect on limb formation¹⁹, suggesting that this gene has possible redundancy with other Hox genes. In this respect, *Hoxa7* is also expressed in the posterior of the forelimb bud and is induced by RA^{20,21}. Moreover, we found that, like *Hoxb8*, *Hoxa7* is expressed in a forelimb-specific fashion (Supplementary Fig. 3). Intriguingly, *Hoxa7* is also a predicted target of *miR-196*, with several conserved matches to the 5' portion of the miRNA known as the 'seed'²². We did not observe changes in *Hoxa7* mRNA in response to *miR-196* misexpression (data

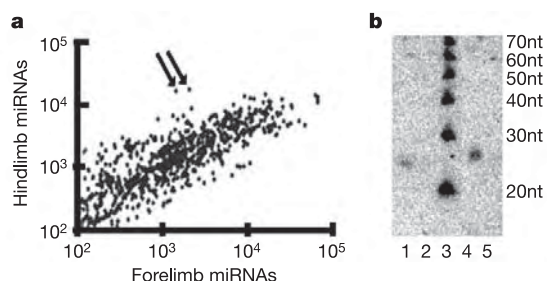


Figure 2 | Hindlimb-specific expression of *miR-196*. **a**, Representation of miRNA array analysis, comparing the expression of individual miRNAs (dots) in E10.5 mouse forelimb and hindlimb buds (in arbitrary units). Abundance of an individual miRNA in the hind- and forelimb is shown by its relative position along the logarithmically scaled y and x axes, respectively. Arrows indicate features corresponding to *miR-196*. **b**, Northern blot hybridization detected *miR-196* in extracts from hindlimbs of E10.5 mouse and stage-22 chick (lanes 1 and 4, respectively) but not in mouse and chick forelimb buds (lanes 2 and 5, respectively). Data are representative of four independent samples. The lengths of DNA oligomers (lane 3) used as size markers are specified next to the blot in nucleotides (nt).

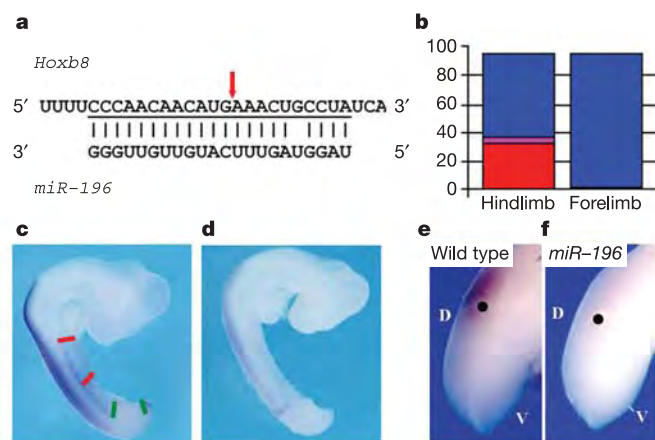


Figure 3 | *miR-196* downregulates *Hoxb8* accumulation. **a**, Sequence of the 3' UTR of *Hoxb8* complements *miR-196*. An arrow indicates the 5' end of the primary cleavage product. **b**, 5' RACE analysis in hindlimb and forelimb. Of the 96 hindlimb clones sequenced, 33 yielded a sequence consistent with *miR-196*-directed cleavage (red); four were also truncated *Hoxb8* clones, but cleavage was outside the miRNA-binding site (pink); and 59 were sequences unrelated to *Hoxb8* (blue). In the forelimb, no clones were consistent with *miR-196*-directed cleavage. **c**, By whole-mount *in situ* hybridization with a *Hoxb8* probe, an expression domain of *Hoxb8* was detected in the forelimb field (red bars), but not in the hindlimb field (green bars), of a stage-16 chick embryo ($n = 8/8$). **d**, Early pan infection with RCAS::miR-196 resulted in downregulation of *Hoxb8* ($n = 6/6$). **e**, An RA-soaked bead implanted into the anterior aspect of a stage-22 wild-type forelimb induced *Hoxb8* expression ($n = 8/10$). **f**, Only marginal induction of *Hoxb8* expression was detected on implantation of an RA-soaked bead in a forelimb infected with RCAS::miR-196 ($n = 6/8$). Anterior view; D, dorsal; V, ventral.

not shown), however, indicating that if *miR-196* is repressing *Hoxa7*, it is reducing *Hoxa7* protein without substantially destabilizing the *Hoxa7* transcript. Such a mechanism would be consistent with the results of a heterologous reporter assay showing that a *Hoxa7* untranslated region (UTR) fragment containing the *miR-196* seed matches predominantly mediates *miR-196*-dependent repression through the reduction of protein rather than mRNA levels¹⁷.

The experiments described here indicate that *miR-196* may be an *in vivo* inhibitor of *Hoxb8* in the hindlimb, and thereby may be responsible for the inability of ectopic RA to induce *Hoxb8* in the hindlimb. Low *Hoxb8* expression and *miR-196*-directed degradation was detected in the naive hindlimb bud by 5' RACE, indicating that *miR-196* activity is a component of *Hoxb8* regulation in the unmanipulated limb. Notably, however, loss of miRNA activity in the *Dicer*-deficient hindlimb did not, in itself, result in *Hoxb8* induction, suggesting that the primary level of regulation of forelimb-specific *Hoxb8* expression is transcriptional and independent of small regulatory RNAs.

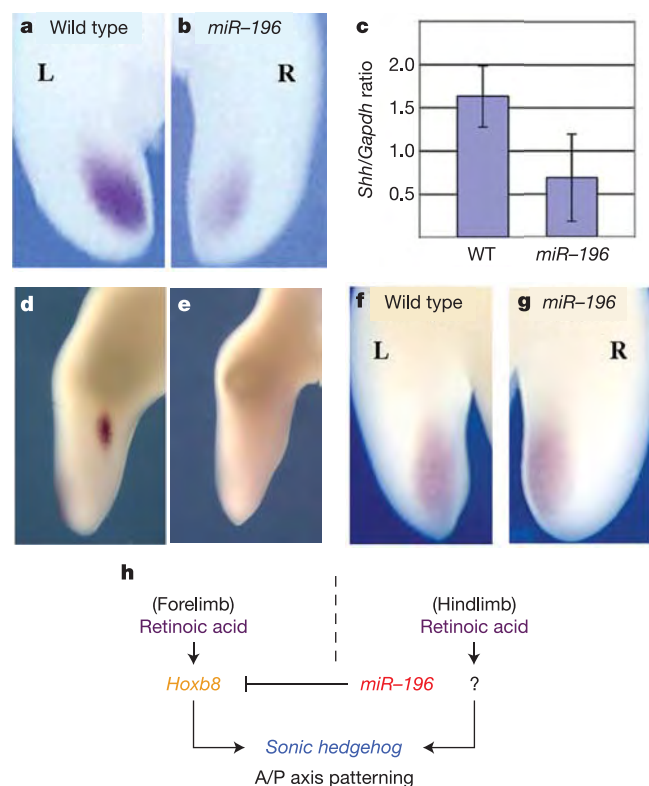


Figure 4 | *miR-196* downregulates *Shh* in the chick forelimb. **a**, Expression of *Shh* in the left (L) limb of stage-23 embryo. Posterior view ($n = 20/20$). **b**, In the right forelimb (R) of the same embryo, endogenous *Shh* expression was diminished 2 d after infection with *RCAS::miR-196* ($n = 18/20$). **c**, Three untreated sample tubes, each containing four forelimbs (stage 23), and three corresponding sample tubes with *RCAS::miR-196*-infected limbs were subjected to real-time PCR quantification of *Shh* mRNA. Three replicate runs were done on each sample tube. Blue bars represent the expression of *Shh*, normalized to *Gapdh*, in wild-type limbs (mean $\pm$ s.d., 1.62 ± 0.35) and limbs infected with *RCAS::miR-196* (0.68 ± 0.51). The difference in the mean value between the *miR-196*-infected sample and the untreated control was significant (one-tailed t -test, $P = 0.029$). **d**, One and a half days after an RA-soaked bead (1 mg ml^{-1}) was implanted into the anterior aspect of stage-20 forelimbs, an ectopic *Shh* expression domain was detected by whole-mount *in situ* hybridization ($n = 6/6$). **e**, *RCAS::miR-196* infection inhibited the ectopic expression of *Shh* in the anterior ($n = 5/8$). **f**, **g**, *Shh* expression was comparable in the left uninfected hindlimb (**f**) and the right *RCAS::miR-196*-infected hindlimb (**g**) of chick embryos ($n = 20/20$). **h**, Model of the epistatic relations among *miR-196*, RA, *Hoxb8* and *Shh* in the developing limbs. A/P, anterior–posterior.

Thus, in normal limb development, the role of *miR-196* seems to be to safeguard against inappropriate Hox activity in the hindlimb. This conclusion fits well with the report that the genes that are downregulated when a miRNA is delivered to human cells are preferentially those that are expressed at low levels in tissues that normally express the miRNA²³. It thus seems that a chief role of some miRNAs in vertebrate development may be to prevent inappropriate activity of genes in domains where they are already repressed transcriptionally. Some miRNAs have been experimentally implicated to have roles in other facets of vertebrate development, including *miR-181* in haematopoiesis²⁴, *miR-430* in brain morphogenesis²⁵ and *miR-1* in heart development²⁶. In contrast to our findings, *miR-1* and its target *hand2* are predominantly expressed in the same cells, enabling *miR-1* to have a key role in regulating the switch between cardiomyocyte differentiation and proliferation²⁶. Together, these two studies indicate that these intriguing regulators of gene activity can take on diverse roles in coordinating vertebrate developmental and physiological processes.

METHODS

Mice and organ culture. Mice were housed and handled in accordance with protocols approved by the Institutional Animal Care and Use Committee of Harvard Medical School. Male mice carrying one copy of the *Prx1::Cre* allele and one *Dicer*^{flxed} allele were crossed to *Dicer*^{flxed/flxed} females. Cre recombinase, driven by the *prx1* enhancer, excises a required region in the RNase IIIb domain to yield a nonfunctional *Dicer* allele in limb buds¹³. Timed-pregnant females were killed at E11.5, embryos were dissected, and limbs were separately cultured in hanging drops. After 12 h of incubation in DMEM medium supplemented with 10% fetal calf serum, penicillin and streptomycin with or without 100 nM all-trans RA (Sigma), limbs were fixed in 4% paraformaldehyde for 4 h and processed for *Hoxb8* *in situ* hybridization.

MicroRNA–cDNA probe and expression array hybridization. Total RNA was isolated from E10.5 mouse fore- and hindlimbs with Trizol (Invitrogen) according to the manufacturer's instructions. Small RNAs were size-fractionated, ligated to adaptor oligonucleotides, reverse-transcribed and amplified. Labelled probes (Cy5 for the hindlimb sample and Cy3 for the forelimb sample) were hybridized to an expression array as described¹⁵. After hybridization, the array was scanned (Genepix pro 4000b; Axon) and analysed. Along with the vertebrate spots on the array, spots for all known *Caenorhabditis elegans* miRNAs are printed, most of which should not be hybridized to a vertebrate probe. Thus, background was set at a score equal to 95% that of the spots from the *C. elegans* section of the array¹⁵.

5' RACE of *Hoxb8*. Total RNA was obtained from a pool of 30 E10.5–11 mouse hind- and forelimbs and was subjected to modified 5' RACE as described¹⁷ with the following primers: 5'-CCATAAGCAATTCACAGATACAGG-3' and 5'-GGTTGCGAGGAAAGATG-3'.

Generation of *RCAS::miR-196*. A 500-bp fragment of genomic DNA surrounding the chicken *miR-196-1* locus (chromosome 27, *HoxB* cluster) was amplified by PCR. An *ApaI* site was appended to the 5' end and an *EcoRI* site was appended to the 3' end by using the following primers (restriction sites are in parentheses): 5'-AATTCC(GGGCCC)CTCTATTTGTCAACTATTGTACG-3' and 5'-G(GAATTC)GCATTTTGGCCTCCGAGAGG-3'. The PCR fragment was then cloned, by means of the *ApaI* and *EcoRI* sites, downstream of the RNA polymerase III U6 promoter, into a pBS-U6 plasmid. The whole U6 promoter and *miR-196* genomic DNA were then excised with *Clal* and cloned into the *RCAS* virus. *RCAS::miR-196* viral particles at a titre of 10^{10} particles per ml were collected from the medium of transfected chicken embryonic fibroblasts. Proper transcription and processing of mature *miR-196-1* was confirmed by northern blots of total RNA extracted from chicken embryonic fibroblasts (data not shown).

Chicken embryo manipulations and *in situ* hybridization. Fertilized eggs were obtained from SPAFAS and incubated at 37 °C, and the embryos were staged according to ref. 27. Eggs were incubated up to stage 7–8 and then the whole embryo was targeted by multiple injections of *RCAS::miR-196*. Alternatively, at stage 12–13 the coelomic cavity was targeted to infect the lateral plate mesoderm. Resin beads were soaked in 100 nM all-trans RA in dimethylsulphoxide for 1 h and then implanted into the anterior of stage-22 chick forelimbs for a further 4 h, as described¹⁰, except that AG-1X8 beads (Bio-Rad) were used. Alternatively, RA-soaked AG-1X2 beads ($1 \text{ mg ml}^{-1} = 300 \text{ mM}$) were implanted into stage-20 limbs that were allowed to develop *in ovo* for 36 h more⁴. Embryos were then collected and fixed in 4% paraformaldehyde overnight. Whole-mount *in situ* hybridization and probes have been described^{4,28}. The *Hoxa7* probe was

amplified directly by PCR from chicken genomic DNA and transcribed, without subcloning, by using the following primers: 5'-ACCTACACCCGCTACCAGAC-3' and 5'-TGTAATACGACTCACTATAGGGCCCTCTTCCTCATCTTCTTCCA-3'. **Quantitative real-time PCR for chick *Shh*.** Three untreated sample tubes, each containing four stage-23 forelimbs, and three corresponding sample tubes with *miR-196*-infected limbs were subject to quantification of *Shh* mRNA. Three replicate runs were done on each sample tube with a Lightcycler 2000 (Roche) using SYBER Green DNA Master Mix (Roche) and the following primers: GAPDH-5', 5'-CGGAGTCAACGATT-3'; GAPDH-3', 5'-ATAACACGCTTACACC-3'; *Shh*-5', 5'-TGCTAGGGATCGGTGGATAG-3'; *Shh*-3', 5'-ACAA GTCAGCCCAGAGGAGA-3'. A 'no RT' control was done in parallel (data not shown). One-tailed *t*-test determined the significance of the difference in the mean value between the *miR-196*-infected sample and the untreated control.

Received 24 June; accepted 10 August 2005.

- Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- Ambros, V. The functions of animal microRNAs. *Nature* **431**, 350–355 (2004).
- Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
- Riddle, R. D., Johnson, R. L., Laufer, E. & Tabin, C. Sonic hedgehog mediates the polarizing activity of the ZPA. *Cell* **75**, 1401–1416 (1993).
- Stratford, T., Horton, C. & Maden, M. Retinoic acid is required for the initiation of outgrowth in the chick limb bud. *Curr. Biol.* **6**, 1124–1133 (1996).
- Helms, J. A., Kim, C. H., Eichele, G. & Thaller, C. Retinoic acid signalling is required during early chick limb development. *Development* **122**, 1385–1394 (1996).
- Lu, H. C., Revelli, J. P., Goering, L., Thaller, C. & Eichele, G. Retinoid signalling is required for the establishment of a ZPA and for the expression of Hoxb-8, a mediator of ZPA formation. *Development* **124**, 1643–1651 (1997).
- Stratford, T., Logan, C., Zile, M. & Maden, M. Abnormal anteroposterior and dorsoventral patterning of the limb bud in the absence of retinoids. *Mech. Dev.* **81**, 115–125 (1999).
- Charite, J., de Graaff, W., Shen, S. & Deschamps, J. Ectopic expression of Hoxb-8 causes duplication of the ZPA in the forelimb and homeotic transformation of axial structures. *Cell* **78**, 589–601 (1994).
- Stratford, T. H., Kostakopoulou, K. & Maden, M. Hoxb-8 has a role in establishing early anterior-posterior polarity in chick forelimb but not hindlimb. *Development* **124**, 4225–4234 (1997).
- He, L. & Hannon, G. J. MicroRNAs: small RNAs with a big role in gene regulation. *Nature Rev. Genet.* **5**, 522–531 (2004).
- Meister, G. & Tuschl, T. Mechanisms of gene silencing by double-stranded RNA. *Nature* **431**, 343–349 (2004).
- Harfe, B. D., McManus, M. T., Mansfield, J. H., Hornstein, E. & Tabin, C. J. The RNaseIII enzyme Dicer is required for morphogenesis but not patterning of the vertebrate limb. *Proc. Natl Acad. Sci. USA* **102**, 10898–10903 (2005).
- Logan, M. *et al.* Expression of Cre Recombinase in the developing mouse limb bud driven by a *Px1* enhancer. *Genesis* **33**, 77–80 (2002).
- Baskerville, S. & Bartel, D. P. Microarray profiling of microRNAs reveals frequent coexpression with neighbouring microRNAs and host genes. *RNA* **11**, 241–247 (2005).
- Mansfield, J. H. *et al.* MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nature Genet.* **36**, 1079–1083 (2004).
- Yekta, S., Shih, I. H. & Bartel, D. P. MicroRNA-directed cleavage of *HOXB8* mRNA. *Science* **304**, 594–596 (2004).
- Llave, C., Xie, Z., Kasschau, K. D. & Carrington, J. C. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. *Science* **297**, 2053–2056 (2002).
- van den Akker, E. *et al.* Axial skeletal patterning in mice lacking all paralogous group 8 Hox genes. *Development* **128**, 1911–1921 (2001).
- Kim, M. H. *et al.* Retinoic acid response element in *HOXA-7* regulatory region affects the rate, not the formation of anterior boundary expression. *Int. J. Dev. Biol.* **46**, 325–328 (2002).
- Min, W. *et al.* 307-bp fragment in *HOXA7* upstream sequence is sufficient for anterior boundary formation. *DNA Cell Biol.* **17**, 293–299 (1998).
- Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P. & Burge, C. B. Prediction of mammalian microRNA targets. *Cell* **115**, 787–798 (2003).
- Lim, L. P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–773 (2005).
- Chen, C. Z., Li, L., Lodish, H. F. & Bartel, D. P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **303**, 83–86 (2004).
- Giraldez, A. J. *et al.* MicroRNAs regulate brain morphogenesis in zebrafish. *Science* **308**, 833–838 (2005).
- Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets *Hand2* during cardiogenesis. *Nature* **436**, 214–220 (2005).
- Hamburger, V. & Hamilton, H. L. A series of normal stages in the development of the chick embryo. *J. Morphol.* **88**, 49–82 (1951).
- Nelson, C. E. *et al.* Analysis of *Hox* gene expression in the chick limb bud. *Development* **122**, 1449–1466 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. McGlinn and S. Nissim for critically reading the manuscript. This work was funded by grants from the NIH (to C.J.T. and to D.P.B.). E.H. was supported by a 'Dorot' fellowship and 'Bikura' award. J.H.M. is supported by a Kirchstein postdoctoral fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.J.T. (tabin@genetics.med.harvard.edu).

Endophilin and CtBP/BARS are not acyl transferases in endocytosis or Golgi fission

Jennifer L. Gallop¹, P. Jonathan G. Butler¹ & Harvey T. McMahon¹

Endophilins have been proposed to have an enzymatic activity (a lysophosphatidic acid acyl transferase or LPAAT activity) that can make phosphatidic acid in membranes^{1–3}. This activity is thought to change the bilayer asymmetry in such a way that negative membrane curvature at the neck of a budding vesicle will be stabilized. An LPAAT activity has also been proposed for CtBP/BARS (carboxy-terminal binding protein/brefeldin A-ribosylated substrate), a transcription co-repressor that is implicated in dynamin-independent endocytosis and fission of the Golgi in mitosis^{4–6}. Here we show that the LPAAT activity associated with endophilin is a contaminant of the purification procedure and can be also found associated with the pleckstrin homology domain of dynamin. Likewise, the LPAAT activity associated with CtBP/BARS is also a co-purification artefact. The proposed locus of activity in endophilins includes the BAR domain, which has no catalytic site but instead senses positive membrane curvature. These data will prompt a re-evaluation of the molecular details of membrane budding.

The lipid composition of a membrane helps to define the identity of an organelle, the flexibility and permeability of the bilayer, and its interaction partners. In phospholipid biosynthesis, glycerol-3-phosphate is acylated to form lysophosphatidic acid (LPA), which is further acylated (on the *sn*-2 position) by LPAAT activity to form phosphatidic acid. Phosphatidic acid is a precursor for glycerophospholipids (including phosphatidylethanolamine and phosphatidylinositols), some of which are increasingly recognized to have important roles in membrane traffic. The highly active LPAATs that are involved in major lipid biosynthetic pathways are transmembrane proteins: LPAAT- α has ubiquitous tissue distribution and is localized to the ER; LPAAT- β has more limited tissue distribution, is highly expressed in various tumour cells and is a drug target^{7,8}.

Low LPAAT activities have been also described for endophilin and CtBP/BARS, cytosolic protein families that are implicated in membrane trafficking^{1,2,4,5}, and have generated great interest. Endophilin A proteins are enriched presynaptically and are required for efficient synaptic vesicle retrieval^{3,9}. These endophilins can tubulate membranes *in vitro* by their N-BAR domain¹⁰ (and our unpublished data) and their intrinsic LPAAT activity is proposed to change membrane curvature by altering the lipid composition^{1,3,11}. Similarly, endophilin B1 is proposed to have LPAAT activity and to be involved in maintaining mitochondrial morphology^{2,12}. Another protein with proposed intrinsic LPAAT activity is CtBP3 (also termed BARS50)⁴, a minor splice variant of CtBP1. There is no sequence homology between either of these protein families and transmembrane LPAATs, and there is no structural homology between CtBP proteins and endophilins^{11,13,14}. CtBP3 is found at the Golgi and in the nucleus and is postulated to be involved in Golgi fission⁵, in fission of vesicles involved in basolateral transport from the Golgi to the plasma membrane⁶, in dynamin-independent endocytosis⁶ and in transcriptional regulation¹⁵. Thus, its precise function is debated. CtBP

proteins have a dehydrogenase domain that binds differentially to NAD⁺/NADH and in transcription its co-repressor activity is proposed to be sensitive to the redox potential¹⁶. The proteins bind to transcription repressors such as E1A and RIP140 and to the tumour suppressor adenomatous polyposis coli (APC) through PxDSL motifs^{13,17}. Mouse knockouts of CtBP1 and CtBP2 show gross developmental defects¹⁸. CtBP2 has an alternative promoter that gives rise to Ribeye, a protein that is enriched in dense bodies at ribbon synapses and is likely to function in synaptic vesicle exocytosis through interactions with active zone proteins^{19,20}. CtBP1 is also present in synaptic terminals¹⁹.

The involvement of endophilin in endocytosis and CtBP/BARS in Golgi remodelling has led to a model in which negative curvature at the neck and indeed membrane fission may be aided by a local change in lipid composition from LPA, an inverted-cone-shaped lipid, to phosphatidic acid, a cone-shaped lipid^{1,4,21,22} (Fig. 1a). The similar mechanisms proposed for these non-homologous proteins led us to investigate further the basis of these reported activities. Doubt has previously been cast on the LPAAT activity of endophilin because it tubulates liposomes in the absence of LPAAT substrates¹⁰; however, the LPAAT activity implicates the protein in negative curvature generation and/or stabilization (Fig. 1b) and the *in vitro* tubulation assays do not address this type of curvature. Thus, a different approach is needed to test whether endophilin or CtBP/BARS do indeed have LPAAT activity.

In a previous study¹, the LPAAT activity of endophilin was localized to a region just larger than the N-BAR domain (amino acids 1–293). The activity was low but this might have been due to the non-optimal presentation of substrate or the absence of activators in the assay. We confirmed the observation of LPAAT activity using full-length rat endophilin A1 (Fig. 1c). We also found that the N-BAR domain (residues 1–247) has this activity (Fig. 1d, e).

We solved the structure of the endophilin A1 BAR domain and found, by comparison with other BAR domains that do not have LPAAT activity, that the only region that is different (and thus could be responsible for the activity) is an insertion in the BAR domain of 22 residues on the membrane interaction face that is not present in the structure (our own unpublished data). One might speculate that this region folds on substrate binding; however, we found that a deletion mutant of the region has LPAAT activity identical to that of the wild-type N-BAR domain (Fig. 1d).

Similarly to previous findings¹, we found LPAAT activity in a supernatant obtained after freeze–thaw cycles and high-speed centrifugation of endophilin-transformed bacteria, and none in that obtained after the same treatment of vector-transformed bacteria (Fig. 1e). Transformation with a construct encoding a control lipid-binding domain, namely the dynamin double pleckstrin homology (PH) domain, also had LPAAT activity in its supernatant (Fig. 1e). LPAAT activity co-purified with the PH domain on glutathione S-transferase (GST) purification (Fig. 1f), but was separated away

¹MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

from the protein by gel filtration (Fig. 1f and Supplementary Fig. 1e, f). Notably, purification of low LPAAT activity was more reproducible with the PH domain than with endophilin. The activity was not always observed with endophilin and seemed to correlate inversely with protein expression (that is, when we had much less protein expression, the GST purification was less effective but there was more LPAAT activity). As endophilin is a membrane-binding and tubulating protein, it may well fragment bacterial membranes or bind to small vesicles, resulting in co-purification of LPAAT activity. We consistently found no LPAAT activity in our endophilin BAR domain purified for crystallography.

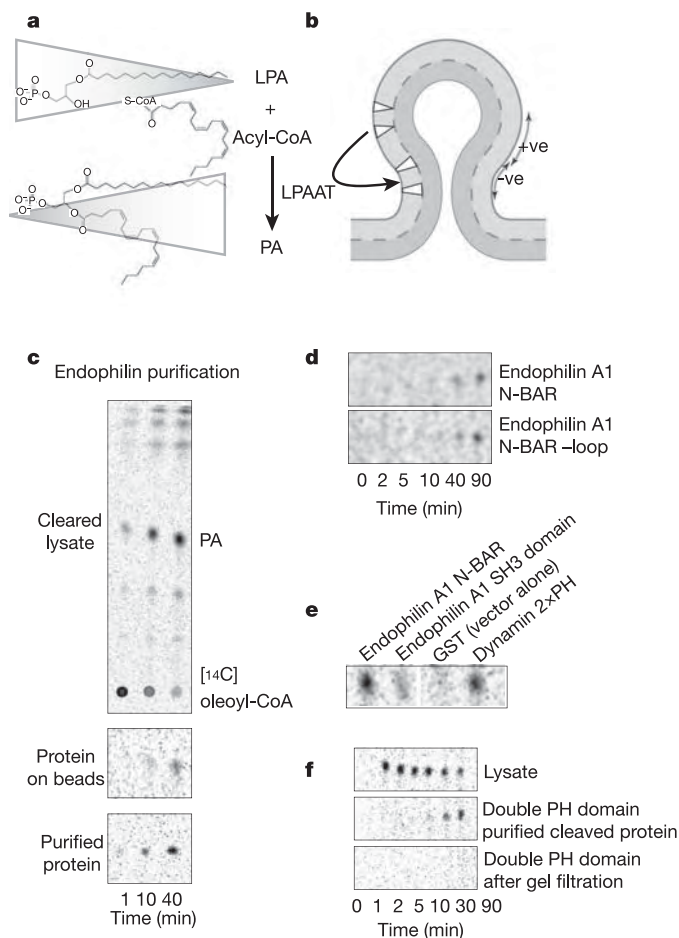


Figure 1 | LPAAT activity may be a co-purification artefact. **a**, LPAAT adds an acyl chain onto LPA to make phosphatidic acid (PA). **b**, LPA and phosphatidic acid should favour different curvatures during vesicle budding. **c**, Endophilin has LPAAT activity, as measured by the production of phosphatidic acid from oleoyl-CoA. Phosphatidic acid was separated by thin layer chromatography and identified by comparison with a standard (see Supplementary Fig. 1). GST-tagged full-length endophilin protein was expressed in BL21 *E. coli* and after affinity purification the GST tag was removed with PreScission. The LPAAT activity of the BL21 lysate was more robust and less variable than that of the purified protein (which was always freshly prepared before measurements). Only the region corresponding to phosphatidic acid is shown. **d**, The N-BAR domain (used at half the protein concentration) also has LPAAT activity, which is not affected by excision of the N-BAR loop. The methods were the same as in **c**. **e**, Although the freeze-thaw/high-speed supernatant from bacteria transformed with vector only does not have LPAAT activity, that from bacteria transformed with the control lipid-binding double PH domain of dynamin does (visible after 90 min of reaction time). **f**, LPAAT assays for purification of the dynamin double PH domain. Activity is present in the lysates (as expected) and also in the purified cleaved protein. No activity is observed after gel filtration. See Supplementary Fig. 1e for purification.

To test the possibility that LPAAT activity is not intrinsic to endophilin, we used a complementation assay in bacteria that has been previously used to identify LPAATs²⁴. A JC201 strain of *Escherichia coli* lacking endogenous LPAAT activity grows at 30 °C but not at 42 °C (ref. 25). This defect at 42 °C is complemented by transformation of an *Escherichia coli* native LPAAT²⁵ (Fig. 2a). Transformation of JC201 with endophilin A1, other BAR domain constructs and control lipid-binding proteins showed that these constructs did not rescue temperature sensitivity (Fig. 2a) or growth kinetics

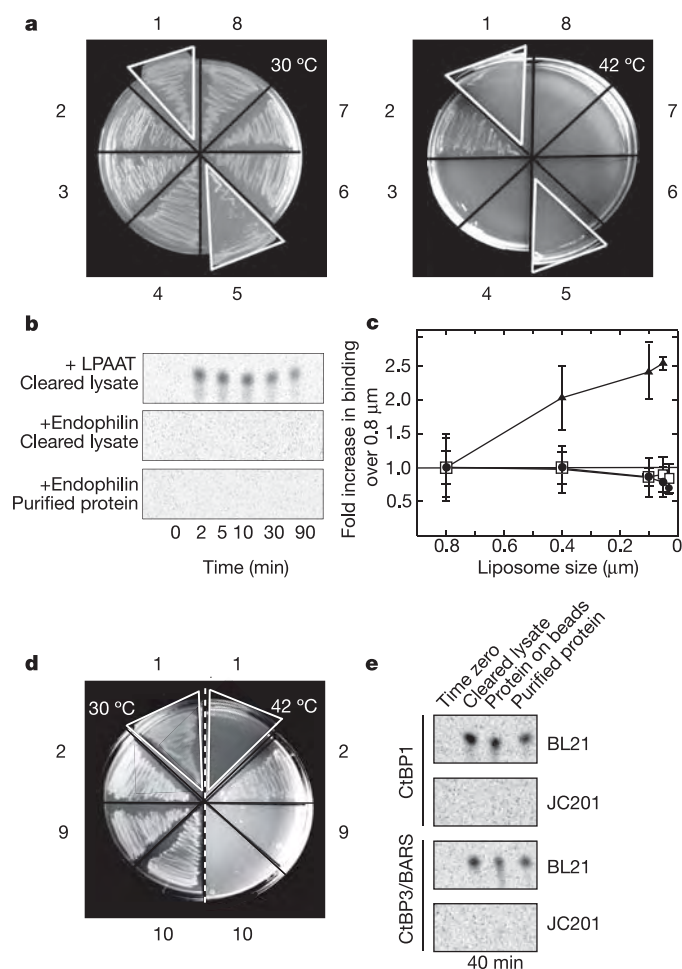


Figure 2 | Endophilin and CtBP/BARS do not have LPAAT activity and endophilin effects positive, rather than negative membrane curvature.

a, *E. coli* that has no endogenous LPAAT activity (JC201 strain) cannot grow at 42 °C (sector 1). This growth defect can be rescued by other LPAATs (sector 2) but not by endophilin (sectors 3–5). LPAAT-negative *E. coli* were transformed with nothing (1); *pPlsC* (the *E. coli* gene encoding LPAAT), which rescues temperature sensitivity (2); GST-tagged full-length endophilin A1 (3); GST-tagged endophilin A1 N-BAR (4); C-terminally His-tagged endophilin A1 N-BAR (5); amphiphysin 1 N-BAR+ (6); epsin 1 ENTH domain (7); and arfaptin 2 (8). **b**, No LPAAT activity was observed for full-length endophilin in purifications from LPAAT-negative JC201 *E. coli*. **c**, The endophilin BAR domain binds better to smaller liposomes (liposomes of increased positive curvature). Filled triangles, endophilin A1 BAR domain; open squares, epsin 1 ENTH domain; filled circles, endophilin A1 N-BAR domain. The N-BAR actively tubulates liposomes and shows no preference in this assay. (See also Supplementary Methods and Supplementary Fig. 4.) Results are the mean $\pm$ s.d. of three experiments. **d**, CtBP1 or CtBP3/BARS does not rescue the temperature-sensitive growth of LPAAT-negative bacteria. JC201 were transformed with nothing (1) as in **a**; *pPlsC* (2) as in **a**; GST-CtBP1 (9); and GST-CtBP3/BARS (10). **e**, LPAAT assays on CtBP/BARS purified from BL21 and JC201 at time zero or after 40 min of incubation with the indicated fractions. CtBP/BARS has no LPAAT activity when purified from JC201.

(Supplementary Fig. 2). Thus, endophilin does not complement LPAAT-defective cells. Endophilin purified from this JC201 strain does not have LPAAT activity (Fig. 2b). In support of the LPAAT activity of endophilin, a previous study showed binding of this protein to the LPAAT substrates palmitoyl-CoA (linked to agarose) and LPA¹. We cast doubt on the significance of these interactions in Supplementary Figs 3 and 4.

Functionally, LPAAT activity was proposed to generate or to aid the formation of negative curvature found at the neck of a budding vesicle. We found that the endophilin BAR domain, as expected from its homology to other BAR domains, actually binds preferentially to smaller liposomes, which have increased positive curvature rather than the negative curvature suggested by the LPAAT hypothesis (Fig. 2c). This indicates that if endophilin does function in formation of the vesicle neck then the curvature generated is positive rather than negative. Notably, the curvature inherent in the BAR domain corresponds more closely to that of the neck than the vesicle dome. The membrane-binding characteristics of the proposed locus of LPAAT activity are therefore at odds with the LPAAT hypothesis and suggest that endophilin has an alternative role in endocytosis.

CtBP/BARS has also been reported to have LPAAT activity⁴. We tested CtBP/BARS in the JC201 complementation assay and again found no rescue of growth at 42 °C (Fig. 2d). CtBP1 and its splice variant CtBP3 were expressed in BL21 and JC201 *E. coli*, and LPAAT activity was tested under the maximum activity conditions reported previously⁴ (Fig. 2e). Although we observed the activity in protein purified from BL21, we did not observe it in protein purified from JC201, despite correct protein folding and the binding of either NADH to the nucleotide binding domain or PxDSL motifs from APC to the substrate binding domain (Supplementary Fig. 5). We further tested the specificity of LPAAT activity of CtBP/BARS using the proposed competitive inhibitor NADH¹⁵; the lack of inhibition observed again supports the idea that the activity is a co-purification artefact (Supplementary Figs 5 and 6).

The dynamic modulation of membrane curvature by localized LPAAT activity is an attractive proposition; however, we have presented evidence that this activity is not contained in endophilin or CtBP/BARS. Indeed, the reported activity of endophilin was less than 1 mol min⁻¹ per mol of endophilin¹, which is much too slow to effect curvature during vesicle invagination unless an activator is used. An involvement of CtBP in membrane trafficking is strongly supported by its localization in dense bodies of synaptic terminals¹⁹, where we propose that it might function in vesicle tethering rather than lipid modification. As in transcriptional regulation, this function is likely to be responsive to the redox balance, which would be a particularly useful readout of cellular energy status given the energy required for neurotransmission.

This study should act as a caution in interpreting enzymatic assays, even those done with purified recombinant protein. Although affinity tags make crude purifications very easy, our results indicate that this approach is not a replacement for rigorous purification. In the absence of evidence for direct and timely conversion of LPA to phosphatidic acid at the vesicle neck, the contribution of fission proteins such as dynamin takes on renewed importance, and the mechanism of negative curvature generation at the neck may well be a relaxation of the bilayer inner and outer leaflet lipid imbalance caused by vesicle neck formation rather than active generation.

METHODS

LPAAT assay. *E. coli* were lysed by French pressing and clarified by ultracentrifugation (186,000g, 40 min) except in Fig. 1e when they were lysed by freeze-thaw and centrifuged (280,000g, 75 min) according to the method of ref. 1. GST protein was bound to glutathione Sepharose (following the instructions of Amersham) and the tag cleaved using thrombin (Serva) or PreScission (Amersham) for 1 h at room temperature in 150 mM NaCl, 20 mM HEPES

pH 7.4, 2 mM DTT. The purification was performed rapidly to avoid inactivation of a possible activity. LPAAT assays were performed in both the above buffer and in 100 mM KCl, 25 mM sucrose, 10 mM Tris pH 8, plus 1 mg ml⁻¹ fatty acid-free BSA and 0.01% Triton X-100 with substrate concentrations (20–200 μM LPA) at 25 °C and 37 °C with ¹⁴C-labelled oleoyl-CoA (Amersham) at a specific activity of 50 or 0.67 μCi ml⁻¹ depending on the final concentration (10–100 μM). Three microlitres of the 100-μl reactions (using 5 and/or 10 μg protein) were spotted onto K60 silica gel chromatography plates (Merck) and run in chloroform:methanol:acetic acid:water, 50:20:10:10:5 followed by autoradiography.

Protein constructs. The following proteins were expressed in bacteria from either pET or pGex vectors: rat endophilin A1 N-BAR domain (residues 1–247), rat endophilin A1 N-BAR-loop domain (deletion of 59–87 and insertion of two glycines), full-length human arfaptin 2, rat epsin 1 ENTH domain (1–164), bovine dynamin 1 double PH domain (510–633), rat amphiphysin 1 N-BAR+ (1–377), full-length mouse CtBP1 and full-length mouse CtBP3 (where the first 13 residues of CtBP1 are replaced with Met,Ser), and GST–(Dros)E-APC (488–597). Proteins were affinity purified by GSH-Sepharose or Ni-NTA agarose (see Supplementary Fig. 1). If a tag is not present then the protein has been previously N-terminally GST tagged and the tag has been cleaved with thrombin.

Curvature sensing. Liposomes made from Folch brain lipid extract at 0.2 mg ml⁻¹ in 20 mM HEPES pH 7.4, 150 mM NaCl, 2 mM DTT were extruded 11 times through cyclopore filters with given pore sizes²⁶. Endophilin N-BAR (2 μM), endophilin BAR (2 μM) and epsin 1 ENTH domain (15 mg) were incubated in a volume of 100 μl for 10 min and sedimented at 100,000g. Sample buffer was added to the pellet and run on SDS–PAGE and stained with Coomassie brilliant blue. The amount of protein bound to the 0.8 μm liposomes was defined as 1 and the fold increase in binding was calculated relative to this value.

Received 20 May; accepted 15 August 2005.

- Schmidt, A. *et al.* Endophilin I mediates synaptic vesicle formation by transfer of arachidonate to lysophosphatidic acid. *Nature* **401**, 133–141 (1999).
- Modregger, J., Schmidt, A. A., Ritter, B., Huttner, W. B. & Plomann, M. Characterization of Endophilin B1b, a brain-specific membrane-associated lysophosphatidic acid acyl transferase with properties distinct from endophilin A1. *J. Biol. Chem.* **278**, 4160–4167 (2003).
- Guichet, A. *et al.* Essential role of endophilin A in synaptic vesicle budding at the *Drosophila* neuromuscular junction. *EMBO J.* **21**, 1661–1672 (2002).
- Weigert, R. *et al.* CtBP/BARS induces fission of Golgi membranes by acylating lysophosphatidic acid. *Nature* **402**, 429–433 (1999).
- Hidalgo Carcedo, C. *et al.* Mitotic Golgi partitioning is driven by the membrane-fissioning protein CtBP3/BARS. *Science* **305**, 93–96 (2004).
- Bonazzi, M. *et al.* CtBP3/BARS drives membrane fission in dynamin-independent transport pathways. *Nature Cell Biol.* **7**, 570–580 (2005).
- Leung, D. W. The structure and functions of human lysophosphatidic acid acyltransferases. *Front. Biosci.* **6**, D944–D953 (2001).
- Umez-Goto, M. *et al.* Lysophosphatidic acid production and action: validated targets in cancer? *J. Cell. Biochem.* **92**, 1115–1140 (2004).
- Schuske, K. R. *et al.* Endophilin is required for synaptic vesicle endocytosis by localizing synaptotagmin. *Neuron* **40**, 749–762 (2003).
- Farsad, K. *et al.* Generation of high curvature membranes mediated by direct endophilin bilayer interactions. *J. Cell Biol.* **155**, 193–200 (2001).
- Huttner, W. B. & Schmidt, A. Lipids, lipid modification and lipid-protein interaction in membrane budding and fission—insights from the roles of endophilin A1 and synaptophysin in synaptic vesicle endocytosis. *Curr. Opin. Neurobiol.* **10**, 543–551 (2000).
- Karbowsky, M., Jeong, S. Y. & Youle, R. J. Endophilin B1 is required for the maintenance of mitochondrial morphology. *J. Cell Biol.* **166**, 1027–1039 (2004).
- Kumar, V. *et al.* Transcription corepressor CtBP is an NAD⁺-regulated dehydrogenase. *Mol. Cell* **10**, 857–869 (2002).
- Nardini, M. *et al.* CtBP/BARS: a dual-function protein involved in transcription co-repression and Golgi membrane fission. *EMBO J.* **22**, 3122–3130 (2003).
- Chinnadurai, G. CtBP family proteins: more than transcriptional corepressors. *BioEssays* **25**, 9–12 (2003).
- Zhang, Q., Piston, D. W. & Goodman, R. H. Regulation of corepressor function by nuclear NADH. *Science* **295**, 1895–1897 (2002).
- Hamada, F. & Bienz, M. The APC tumour suppressor binds to C-terminal binding protein to divert nuclear β-catenin from TCF. *Dev. Cell* **7**, 677–685 (2004).
- Hildebrand, J. D. & Soriano, P. Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. *Mol. Cell. Biol.* **22**, 5296–5307 (2002).
- tom Dieck, S. *et al.* Molecular dissection of the photoreceptor ribbon synapse: physical interaction of Bassoon and RIBEYE is essential for the assembly of the ribbon complex. *J. Cell Biol.* **168**, 825–836 (2005).

20. Schmitz, F., Königstorfer, A. & Sudhof, T. C. RIBEYE, a component of synaptic ribbons: a protein's journey through evolution provides insight into synaptic ribbon function. *Neuron* **28**, 857–872 (2000).
21. Kooijman, E. E. *et al.* Spontaneous curvature of phosphatidic acid and lysophosphatidic acid. *Biochemistry* **44**, 2097–2102 (2005).
22. Shemesh, T., Luini, A., Malhotra, V., Burger, K. N. & Kozlov, M. M. Prefission constriction of Golgi tubular carriers driven by local lipid metabolism: a theoretical model. *Biophys. J.* **85**, 3813–3827 (2003).
23. Brown, A. P., Coleman, J., Tommey, A. M., Watson, M. D. & Slabas, A. R. Isolation and characterisation of a maize cDNA that complements a 1-acyl-*sn*-glycerol-3-phosphate acyltransferase mutant of *Escherichia coli* and encodes a protein which has similarities to other acyltransferases. *Plant Mol. Biol.* **26**, 211–223 (1994).
24. Coleman, J. Characterization of *Escherichia coli* cells deficient in 1-acyl-*sn*-glycerol-3-phosphate acyltransferase activity. *J. Biol. Chem.* **265**, 17215–17221 (1990).
25. Coleman, J. Characterization of the *Escherichia coli* gene for 1-acyl-*sn*-glycerol-3-phosphate acyltransferase (plsC). *Mol. Gen. Genet.* **232**, 295–303 (1992).
26. Peter, B. J. *et al.* BAR domains as sensors of membrane curvature: the amphiphysin BAR structure. *Science* **303**, 495–499 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank members of the laboratory for support and encouragement; A Brown for JC201 and pPlsC; F. Hamada and M. Bienz for GST-APC(two-motif repeats). J.L.G. was the recipient of a Medical Research Council Pre-doctoral Fellowship.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.T.McM. (hmm@mrc-lmb.cam.ac.uk).

Global analysis of protein phosphorylation in yeast

Jason Ptacek^{1*}, Geeta Devgan^{2*}, Gregory Michaud^{3*}, Heng Zhu², Xiaowei Zhu², Joseph Fasolo², Hong Guo³, Ghil Jona², Ashton Breitkreutz⁴, Richelle Sopko⁴, Rhonda R. McCartney⁵, Martin C. Schmidt⁵, Najma Rachidi⁶, Soo-Jung Lee⁷, Angie S. Mah⁸, Lihao Meng³, Michael J. R. Stark⁶, David F. Stern⁷, Claudio De Virgilio⁹, Mike Tyers⁴, Brenda Andrews^{4,10}, Mark Gerstein¹, Barry Schweitzer³, Paul F. Predki³ & Michael Snyder^{1,2}

Protein phosphorylation is estimated to affect 30% of the proteome and is a major regulatory mechanism that controls many basic cellular processes^{1–3}. Until recently, our biochemical understanding of protein phosphorylation on a global scale has been extremely limited; only one half of the yeast kinases have known *in vivo* substrates and the phosphorylating kinase is known for less than 160 phosphoproteins. Here we describe, with the use of proteome chip technology⁴, the *in vitro* substrates recognized by most yeast protein kinases⁵: we identified over 4,000 phosphorylation events involving 1,325 different proteins. These substrates represent a broad spectrum of different biochemical functions and cellular roles. Distinct sets of substrates were recognized by each protein kinase, including closely related kinases of the protein kinase A family and four cyclin-dependent kinases that vary only in their cyclin subunits. Although many substrates reside in the same cellular compartment or belong to the same functional category as their phosphorylating kinase, many others do not, indicating possible new roles for several kinases. Furthermore, integration of the phosphorylation results with protein–protein interaction^{6–10} and transcription factor binding data^{11,12} revealed novel regulatory modules. Our phosphorylation results have been assembled into a first-generation phosphorylation map for yeast. Because many yeast proteins and pathways are conserved, these results will provide insights into the mechanisms and roles of protein phosphorylation in many eukaryotes.

To develop a kinase–substrate map for eukaryotes, we determined the substrates recognized by 87 different yeast protein kinases and bovine protein kinase A, by using a yeast proteome array and the scheme depicted in Fig. 1a. A total of 82 unique kinases representing most of the 122 yeast protein kinases⁵ were tested; two cyclin-dependent kinases, Pho85 (in complex with Pcl1, Pcl2, Pcl9 and Pho80) and Cdc28 (in complex with Cln2 and Clb5), were also analysed.

Each kinase was incubated separately with two yeast proteome microarrays in the presence of [γ -³³P]ATP (Fig. 1b). The microarrays contained about 4,400 proteins spotted in duplicate on the array. The arrays also contained a variety of control proteins including three protein kinases that served both as positive controls and as landmarks for the identification of phosphorylation signals. For each experiment, two slides were also incubated in the absence of a protein kinase serving to identify protein kinases on the array that autophosphorylate. Four protein kinases (Rim15, Dbf2, Hsl1 and Rad53) that contained inactivating mutations in their catalytic domain were used

as negative controls and exhibited signals identical to those obtained in the absence of protein kinase. The extent of phosphorylation was measured with algorithms specifically designed to detect positive signals. Proteins that were reproducibly phosphorylated in the presence of active kinase relative to the control slides were scored as positive substrates. All results are accessible at <http://networks.gersteinlab.org/phosphorylome/>.

Approximately 4,200 phosphorylation events affecting 1,325 proteins were identified from the 87 yeast protein kinase assays. Each kinase recognized between 1 and 256 substrates with an average of 47 substrates per kinase. A distinct set of substrates was phosphorylated by each protein kinase, indicating that every kinase has a unique substrate recognition profile. Most (73%) substrates were recognized by fewer than three kinases, indicating a strong preference of particular kinases for specific substrates (Supplementary Fig. 1). The largest class of proteins phosphorylated by the protein kinases was transcription factors (311 phosphorylations; $P < 10^{-99}$).

Inspection of the substrate list revealed that at least 14 known *in vivo* substrates of particular kinases were identified (Supplementary Table 1). In addition, each kinase phosphorylated proteins residing in the same cellular location and/or functional category as the kinase. For example, Ark1 (actin-regulating kinase) phosphorylated three substrates involved in late secretory functions, a known role for actin; two of these, Sla1 and Ent2, reside at the cell cortex, the same location as actin.

To determine whether other kinase–substrate pairs represent *in vivo* phosphorylation events, we tested whether the phosphorylation of several candidate substrates depended on the identified kinase *in vivo*. Substrates of six kinases were assayed for loss of phosphorylation by either a reduction or an absence of a signal, a mobility shift, or both in kinase deletion strains relative to wild type (Fig. 1c, i–vi). Differences were observed in 12 cases. Interestingly, in at least five cases we observed that strains lacking the kinase gene had significantly altered levels of the putative substrate (Fig. 1c, iii–vi); in three cases kinase loss increased substrate levels, and in two the substrate levels decreased, indicating that the kinases control protein levels directly or indirectly. Overall, 9.2% of substrates exhibited a reduced phosphorylation, a mobility shift or a markedly altered level of substrate. This is presumably a significant underestimate of the number of *in vivo* substrates for these kinases because many proteins do not exhibit mobility shifts upon phosphorylation, may be modified by redundant kinases or have multiple phosphorylations that mask the loss of one or more phosphorylations. Nonetheless, our *in vivo*

¹Department of Molecular Biophysics and Biochemistry, and ²Molecular, Cellular and Developmental Biology, Yale University, New Haven, Connecticut 06511, USA. ³Invitrogen Corporation, 1600 Faraday Avenue, Carlsbad, California 92008, USA. ⁴Department of Medical Genetics and Microbiology, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada. ⁵Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261, USA. ⁶Division of Gene Regulation and Expression, School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK. ⁷Department of Pathology, Yale University School of Medicine, PO Box 208023, 310 Cedar Street BML 342, New Haven, Connecticut 06520-8023, USA. ⁸Department of Biology, California Institute of Technology, Pasadena, California 91125, USA. ⁹Department of Microbiology and Molecular Medicine, CMU, University of Geneva, 1211 Geneva, Switzerland. ¹⁰Banting & Best Department of Medical Research, University of Toronto, Room 4285, Medical Sciences Building, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada.

*These authors contributed equally to this work.

validation results indicate that the proteome microarray approach has identified many genuine substrates of protein kinases and that phosphorylation can markedly affect protein levels.

We also examined substrate profiles in closely related kinases, a

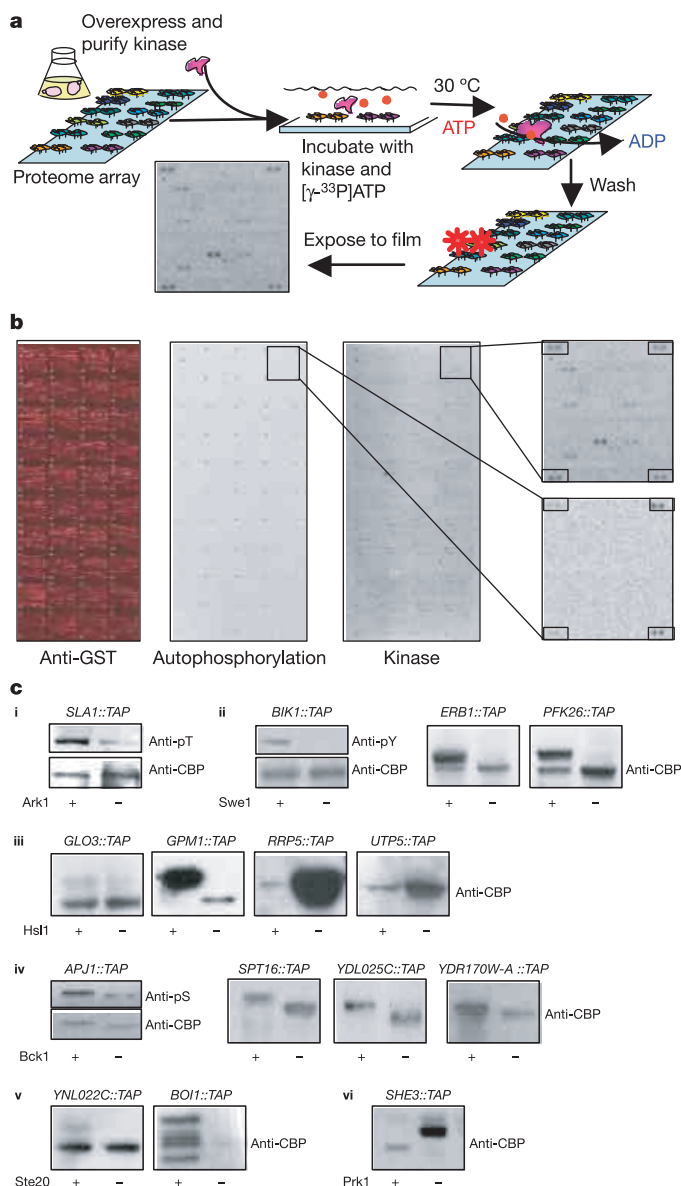


Figure 1 | Identification of kinase substrates using protein chips followed by *in vivo* validation. **a**, Overall scheme to identify kinase substrates. Each kinase was overexpressed, purified and assayed on protein chips containing about 4,400 proteins spotted in duplicate. **b**, Kinase assays on protein chips. Two protein chips were used for every kinase assayed. In addition, two protein chips were probed in the absence of kinase to identify proteins on the chip that autophosphorylate. Commercial kinases were spotted at many defined locations, shown in the four corners of the two boxes on the right; these kinases autophosphorylated in our assay and served as landmarks for the identification of phosphorylation signals. The slide on the left is a representative slide probed with anti-GST antibodies indicating the amount of fusion protein present on the proteome slide. **c**, *In vivo* validation of targets identified on the proteome microarray: *ARK1* (i), *SWE1* (ii), *HSL1* (iii), *BCK1* (iv), *STE20* (v) and *PRK1* (vi) were deleted from the TAP-tagged strains indicated. From the kinase-deleted strains, the tagged proteins were purified and their phosphorylation status compared with wild-type tagged proteins. Immunoblots were probed with anti-phosphotyrosine antibody (i, top panel), anti-phosphotyrosine antibody (ii, top panel) or anti-phosphoserine antibody (iv, top panel). In addition, protein isoforms and protein levels were monitored with anti-CBP antibody (i–vi).

common feature of eukaryotes. Yeast protein kinase A homologues Tpk1 and Tpk3 are 84% identical in amino-acid sequence and 67% and 76% identical to Tpk2, respectively. Strains lacking all three are nonviable, whereas those containing any one of the three Tpk propagate, indicating that each is genetically redundant for cell growth¹³. To determine whether the Tpk kinases are functionally redundant biochemically, each Tpk was directly tested for reactivity with substrates by using proteome arrays prepared and probed at the same time. As shown in Fig. 2, Tpk1, Tpk2 and Tpk3 recognized 256, 29 and 79 substrates, respectively; however, only 8 were recognized by all three kinases and 39 were recognized by two of the three. The vast majority (87.7%) were recognized by only one of the Tpk, indicating that each kinase has a unique substrate specificity; 86.6% of the bovine cAMP-dependent protein kinase (PKA) targets were also substrates of Tpk1. In comparison, two slides probed by the same Tpk showed a greater than 90% substrate overlap. Thus, the closely related Tpk have distinct substrate specificities. These results are consistent with the observation that Tpk1 and Tpk3, although redundant with Tpk2 for cell growth, have different roles in pseudo-hyphal growth¹⁴.

In addition to analysing the Tpk, we examined whether a protein kinase complexed with different regulators recognizes similar or distinct sets of substrates. The substrates recognized by the cyclin-dependent kinase Pho85, purified by itself or complexed with either Pcl1, Pcl2, Pcl9 or Pho80 cyclins, were determined in parallel. Pho85 purified alone recognized only 12 targets, indicating that it has weak activity, as expected. However, between 60 and 255 substrates were observed in the presence of different cyclins. Pho85 complexed with different cyclins exhibited various degrees of overlap in the substrates phosphorylated *in vitro* (Fig. 2). Nearly half (29/60) of the Pho85–Pcl2 substrates overlap with the 89 substrates of Pho85–Pcl9, and most (48/89) of the Pho85–Pcl9 substrates overlap with the 255 substrates of Pho85–Pho80. Pho85–Pcl1 also shares a high degree (43.2%) of substrate preference with Pho85–Pho80, but exhibits very little overlap with Pho85–Pcl9.

These different results indicate that the amino-acid differences of the Tpk and Pho85 cyclins have a considerable influence in substrate recognition. These studies further provide a molecular explanation of why eukaryotic cells have multiple protein kinases with a high degree of sequence similarity: each has different biochemical propensities for particular substrates.

The substrates phosphorylated by the different kinases were also searched for common sequence motifs¹⁵. Consensus motifs were identified for 11 kinases; these are similar to the sequence motifs determined for kinase orthologues in other species (Table 1).



Figure 2 | Comparison of substrates targeted by related kinases.

a, Comparison of the substrate of the different Tpk. Each Tpk kinase has a unique substrate recognition profile; 86.6% of PKA targets are also Tpk1 substrates. **b**, Comparison of the substrates recognized by the different Pho85–cyclin complexes. Pho85, when assayed in the presence of different cofactors, displays different specificities, indicating that the cyclin subunits have a significant impact on substrate recognition.

Table 1 | Summary of motif results

Kinase	Pratt pattern	P	Sites (proteins)	Hits/total substrates	Total protein with sites	Published pattern/site
ARK1	L-x(4)-T-x-[GL]-x-[ST]	4.29×10^{-7}	7 (6)	6/8	209	[LI]-x(2)-Q-x-T-G
CDC28	T-P	2.33×10^{-6}	98 (38)	39/43	2,336	[ST]-P
CKA1	T-x(2)-D	1.04×10^{-3}	42 (18)	18/19	2,487	[ST]-x(2)-[DE]
CMK2	R-x(2)-[ST]-x-[ST]	5.42×10^{-3}	7 (7)	7/9	1,274	R-x(2)-[ST]
DUN1	[ST]-x(3)-S-S	8.63×10^{-5}	33 (15)	15/18	1,530	GSSAS*AS*AS*SLEM (SML1 site)
HRR25	S-x(2)-S	4.75×10^{-2}	58 (14)	14/14	3,288	S-x(2)-[ST]
PKA	R-[KR]-x-S	5.57×10^{-40}	96 (55)	55/56	759	R-[KR]-x-S
PRK1	L-x(4)-[ST]	2.24×10^{-1}	259 (40)	40/41	3,810	[LI]-x(2)-Q-x-T-G
RIM11	S-x(3)-[ST]-x(5)-S-x-[ST]	5.25×10^{-10}	25 (13)	13/17	454	S-x(3)-S
SKY1	D-x(5)-S	8.28×10^{-6}	153 (39)	39/40	2,840	YRTRDAPRERS*PTR (NPL3 site)
YCK2	S-x(2)-D	1.26×10^{-11}	715 (197)	197/220	2,907	S-x(2)-[DE]

Consensus sites were found for 11 protein kinases. These are similar to those reported in the literature. Also summarized are the number of consensus sites in the target proteins (sites (proteins)), the fraction of identified substrates with a consensus site (hits/total substrates) and the total number of proteins on the array with consensus sites (total protein with sites).

Although many of the substrates contain the consensus phosphorylation sites, many yeast proteins with consensus sequences are not kinase targets. For example, 209 and 3,288 proteins on the array have Ark1 and Hrr25 consensus sites, yet only 8 and 14 proteins were recognized, respectively. Therefore, either additional sequences on the substrate help direct substrate recognition or the consensus phosphorylation site is not accessible to the kinase due to spatial and temporal restrictions. Nevertheless, these studies show the importance of directly assaying for protein phosphorylation with experimental tests.

The 4,200 different protein-kinase–substrate phosphorylations have been assembled into an *in vitro* phosphorylation network (Fig. 3). In many cases the identification of substrates helps to define the role of the kinase in yeast signalling networks more accurately. For example, the phosphorylation of Sla1 and She3 by Ark1 and Prk1, respectively, further explains the role of these kinases in actin

regulation. Forty-nine kinase–substrate interactions are also present in the protein–protein interaction network. Presumably the low overlap between these data sets is because kinase–substrate interactions are expected to be transient with low binding affinities and not detectable by most protein–protein interaction assays.

Many of these interactions may represent *in vitro* targets that do not occur *in vivo*. Filtering the phosphorylation network to contain only the interactions in which kinase and substrate are present in the same cellular compartment or in the same functional categories results in 1,384 (33%; $P < 10^{-99}$) and 768 (18.4%; $P < 10^{-99}$) interactions, respectively (Fig. 3, Supplementary Table 2). Furthermore, of the 29 kinases with functional assignments and 9 or more substrates, 14 showed enrichment of substrates in the same categories as those of the protein kinase (Supplementary Table 3). Filtering is likely to enrich the data set for interactions that occur *in vivo*. However, this approach may also increase the number of false

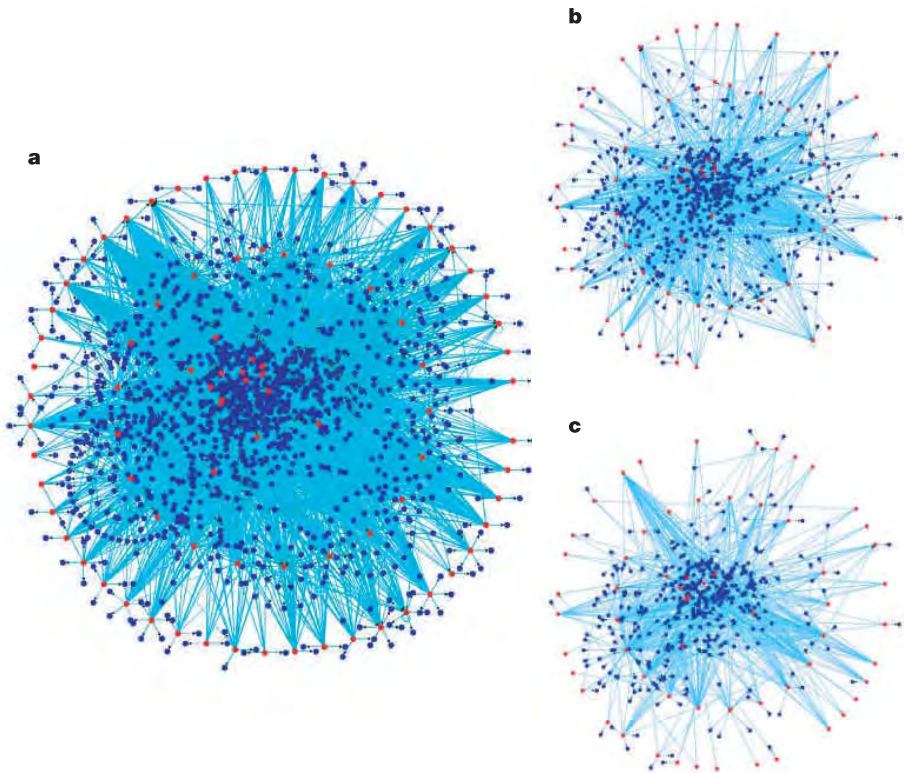


Figure 3 | An *in vitro* phosphorylation map of yeast. **a**, A map showing the connections between kinases and substrates. In all, 87 different kinases (red dots) and 1,325 substrates (blue dots) are represented in the map. **b**, Global localization data can be used to identify only those

phosphorylation events occurring between proteins of the same cellular compartment. **c**, Functional data can be used to identify substrates with similar functions to those of the kinases phosphorylating them.

negatives because the protein localization data may be incomplete and/or inaccurate. Furthermore, the functions of many protein kinases and substrates are either limited or unknown, and thus many substrates would not be expected to have functions assigned to the particular kinase. For example, two kinases involved in a G2/M cell-cycle control pathway, Hsl1 (ref. 16) and Swe1 (ref. 17), phosphorylate proteins involved in glucose metabolism, Gpm1 and Pfk26, respectively; these interactions were validated *in vivo*. Hsl1 and Swe1 localize to the septin ring and nucleus/spindle pole/bud neck, respectively, whereas Gpm1 and Pfk26 are cytosolic proteins involved in glucose metabolism. Presumably, these proteins interact in the cytoplasm and link cytoskeletal function and checkpoint control with glycolysis. The assignment of new functions to protein kinases probably reveals how different cellular pathways can be coordinated by a single regulator.

To understand how phosphorylation may be integrated into global regulatory networks, we also combined the phosphorylation data with transcription factor binding and protein interaction data and generated the first integrated regulatory network for yeast. We then searched this network for common regulatory modules^{6–12,18}. Eight modules were observed, and six (modules 1–6) were of high statistical significance (Fig. 4). At least four modules (1, 2, 3 and 7) have been validated in our studies or from the literature. All of the modules involve kinase–substrate pairs, which we refer to as ‘kinates’ (kinase–substrate pairs). The modules are: (1) interacting kinates, (2) scaffolds, (3) kinase cascades, (4) transcription-factor-regulated kinates, (5) kinate regulon, (6, 7) feedback loops and (8) hetero-substrate regulation. Examples are shown in Fig. 4 and a comprehensive list is available at <http://networks.gersteinlab.org/phosphorylome/>. These results show the utility of integrating different data types; many potential novel regulatory networks, not evident from single data sets, have been identified.

This study is, to our knowledge, the first global investigation of protein phosphorylation by protein kinases using an unbiased approach. A total of 181 substrates of Cdc28–Clb2 were identified previously¹⁹ by computationally searching for substrates with multiple cyclin-dependent kinase sites and assaying for phosphorylation of the glutathione S-transferase (GST)-tagged proteins in cell extracts with a Cdc28-*as* allele. Of 150 tested 24% were preferentially phosphorylated by a Clb28–Clb5 kinase²⁰. We found overlap (10 and 7; $P < 10^{-6}$) between their Cdc28–Clb2 and Cdc28–Clb5 target lists and the 43 substrates we identified with Cdc28–Clb5. Differences between the lists might be due to one of the following: first, the sensitivity of the respective assays; second, an ‘*as* allele’ as opposed to a wild-type kinase; third, the use of different cyclins in the initial screen; fourth, the fact that their assays were performed in cellular extracts and ours were not; and/or fifth, their use of a biased approach in comparison with our survey of most yeast proteins.

We also compared our substrate list with two other studies that mapped *in vivo* phosphorylation sites. The first² mapped phosphorylation sites on 98 proteins that were on our protein chips; our study provides candidate kinases for 50 of these proteins (51%). The second²¹ identified 89 phosphorylation sites that are induced when yeast cells are treated with mating pheromone. Five proteins with these sites are substrates of Pho85–Pcl1, two are substrates for Kss1 and a single substrate was identified for Ste20 and Fus3. The remaining proteins whose phosphorylation is induced by treatment with pheromone may be targets either of additional kinases induced by the pheromone response (for example, Ste7 or Ste11) or of other kinases. Nonetheless, combining the data of our study with *in vivo* data from others provides strong candidates for the kinases that phosphorylate each substrate.

Proteome chips offer many advantages for studying protein phosphorylation. Thousands of proteins can be rapidly screened for enzyme–substrate relationships in an unbiased fashion with very small amounts of reagents and under a variety of test conditions. In addition, closely related kinases with known redundant functions

can be readily differentiated at the molecular level on the basis of their substrate profiles.

Although we were able to identify many known substrates of protein kinases, two-thirds of reported phosphorylations were not observed. This may be due to the absence of 30% of yeast proteins on the array. In addition, the substrates may not be present in sufficient quantity for substrate phosphorylation to be observed. Alternatively, although purified from yeast, the kinases and/or substrates may lack functional adaptors, scaffold proteins or modifications. In principle,

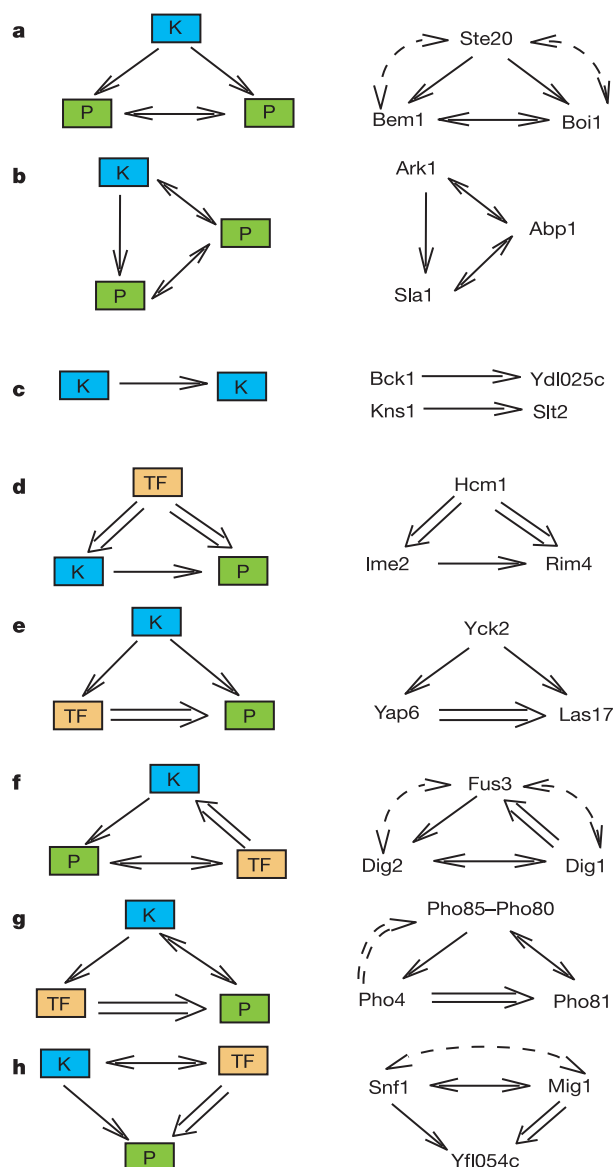


Figure 4 | Integration of other biological data reveals regulatory modules. When protein interaction, transcription factor binding and expression data are considered, many modules within the phosphorylation data are identified. Shown are protein–protein interactions ($\leftrightarrow$), kinase phosphorylations ($\rightarrow$) and transcription factor (TF) regulation ($\Rightarrow$). K, kinase; P, protein. In the following list the modules are numbered from 1 to 8 with their common name in parentheses; also listed are the numbers of occurrences (n) and the statistical significances of such events. **a**, Module 1 (interacting kinates module), $n = 1,563$, $P < 10^{-99}$; **b**, module 2 (scaffold module), $n = 2,448$, $P < 10^{-99}$; **c**, module 3 (kinase cascade module), $n = 147$, $P < 10^{-99}$; **d**, module 4 (TF-regulated kinate module), $n = 145$, $P < 10^{-99}$; **e**, module 5 (kinate regulon module), $n = 92$, $P < 10^{-99}$; **f**, module 6 (kinate feedback loop I module), $n = 25$, $P = 10^{-3}$; **g**, module 7 (kinate feedback loop II module), $n = 11$, not enriched; **h**, module 8 (heterosubstrate regulation module), $n = 14$, not enriched.

the kinase or substrates can be purified from cells grown under different conditions to provide proper modifications on the kinase or substrates. In addition, adaptor proteins (such as cyclins), when identified, can be added to the reactions.

Our assays measure substrate specificity directly at the level of kinase–substrate interaction, which is highly selective because a discrete set of substrates are recognized by each kinase. Nonetheless, phosphorylations that do not normally occur *in vivo* may be identified from this assay. These false positives may be due to either *in vitro* phosphorylation of proteins by kinases that normally reside in other cellular compartments and/or are expressed at different times, or through the absence of adaptor proteins that limit the kinase–substrate interactions. It is unlikely that many false positives are observed by co-purification of a substrate with the intended protein on the array; 80% of all substrates identified can be validated by solution-based assays for substrate phosphorylation and the remainder do not reveal phosphorylation of a co-purifying protein (Supplementary Fig. 2). Combining our data with other information provides a useful method of detecting interactions likely to occur *in vivo*. Because many kinase signalling pathways are highly conserved from fungi to humans¹, our comprehensive identification of the phosphorylation regulatory network in yeast will not only serve as valuable resource for yeast research but will also provide much insight into this important regulatory network in all eukaryotes.

METHODS

Protein purification and proteome arrays. Yeast proteome chips were prepared that were similar to that described previously with the use of yeast strains that overexpressed yeast proteins as GST fusions⁴. About 4,400 yeast strains that consistently express protein of the correct size were used to prepare protein chips (Supplementary Information). Proteins were spotted in duplicate on surface-modified microscope slides by using a 48-pin contact printer (Genomic Solutions). The protein arrays were manufactured at Invitrogen (Branford, Connecticut). Protein kinases containing inactivating mutations in their catalytic domains were prepared by site-directed mutagenesis to mutate an absolutely conserved catalytic residue, Asp, to Ala (ref. 22).

The Pho85 kinases were purified from insect cells²³; the remaining 81 kinases were purified from yeast. Yeast protein kinases were expressed as GST fusions and purified as described previously⁵. Cells were grown in 50–500-ml cultures, harvested and lysed with glass beads in lysis buffer (100 mM Tris-HCl pH 7.4, 100 mM NaCl, 1 mM EGTA, 0.1% 2-mercaptoethanol, 0.1% Triton X-100, protease cocktail (Roche), 1 mM EDTA, 50 mM NaF, 10 mM sodium β -glycerophosphate, 1 mM Na₃VO₄). Kinases were bound to glutathione beads and eluted into kinase buffer (100 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM MgCl₂, 20 mM glutathione, 20% glycerol). Although attempts were made to purify more than 110 protein kinases, only those that were highly active in *in vitro* assays were tested on the protein chips.

Kinase assay and data acquisition. Each kinase was assayed to determine the concentration needed to achieve optimal signal:noise ratio using test protein chips (see Supplementary Information). Proteome arrays were blocked in Superblock (Pierce) with 0.1% Triton X-100 for 1 h at 4°C and probed in duplicate for every kinase. Optimized conditions (see Supplementary Information) consisted of diluting the kinase into kinase buffer plus 0.5 mg ml⁻¹ bovine serum albumin, 0.1% Triton X-100 and 2 μ l of [γ -³³P]ATP (33.3 nM final concentration). Each kinase in buffer was overlaid on two arrays, covered with a coverslip and placed in a humidified chamber at 30°C for 1 h. The slides were washed twice with 10 mM Tris-HCl pH 7.4, 0.5% SDS and once with doubly distilled water before being spun dry and exposed to X-ray film (Kodak). For each experiment, two additional arrays were incubated with kinase buffer in the absence of kinase, which served as autophosphorylation reference slides. Three sets of exposures were taken for each kinase assayed: 1, 3 and 7 days. The X-ray film was scanned at 1,800 dots per inch and each kinase was analysed with GenePix 3.0. The optimal exposure was selected for each kinase and compared with the corresponding autophosphorylation slides.

Data analysis. A computer algorithm was written to identify substrates (see Supplementary Information). Substrates were identified that were two standard deviations above background for at least three of the four protein spots from the two slides assayed for a given kinase. Analysis from the autophosphorylation slides was used to remove proteins showing autophosphorylation. Positive signals were also inspected visually to ensure that each spot was not caused by the presence of an artefact.

A network was generated by combining the substrates for all kinases assayed by using Osprey version 1.2.0 (ref. 24). For each kinase, the substrates were filtered on the basis of functional data from the Munich Information Center for Protein Sequences (MIPS)¹⁸ and localization data in ref. 25. Functional enrichment was performed for the substrates of all the kinases. Substrates enriched with $P < 0.05$ were considered enriched and were then compared with the functional annotation for the kinase. Alignment of the substrates of the Tpk and Pho85 isoforms was performed to obtain a substrate profile for each of the kinases. Modules were identified between transcription factors, kinases and substrates. The requirement for the three-element module was that it should contain at least one phosphorylation interaction and have a total of at least three interactions with all proteins having two interactions. For module 3, kinase–kinase interactions were searched. Common phosphorylation motifs from sets of substrates were identified with a Pratt algorithm¹⁵.

In vivo substrate validation. The phosphorylating kinase was deleted for potential substrates of Ark1, Swe1, Hsl1, Bck1, Prk1 and Ste20 by using the available chromosomally tandem affinity purification (TAP-tagged) strains (see Supplementary Information)^{26,27}. Verified strains were grown and analysed in parallel with the corresponding wild-type TAP-tag strain. TAP-tagged proteins were purified with lysis buffer as above and isolated from the lysates with the use of IgG beads. The bound IgG beads were washed with lysis buffer containing 250 mM NaCl. The beads were heated to 70°C in the presence of NuPAGE loading buffer and eluates were analysed on 10% NuPAGE gels (Invitrogen). Immunoblots were prepared and probed with anti-phosphoserine, anti-phosphothreonine (Qiagen) and anti-phosphotyrosine (Upstate) antibodies to detect loss of phosphorylation. Immunoblots were also probed with anti-calmodulin-binding-peptide antibody (Upstate), which recognizes a portion of the TAP tag, to identify gel mobility shifts. All validated substrates were tested at least twice, and for Glo3, She3 and Gpm1 independent transformants were tested.

Received 21 June; accepted 1 September 2005.

- Manning, G., Plowman, G. D., Hunter, T. & Sudarsanam, S. Evolution of protein kinase signalling from yeast to man. *Trends Biochem. Sci.* **27**, 514–520 (2002).
- Ficarro, S. B. *et al.* Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nature Biotechnol.* **20**, 301–305 (2002).
- Cohen, P. The regulation of protein function by multisite phosphorylation—a 25 year update. *Trends Biochem. Sci.* **25**, 596–601 (2000).
- Zhu, H. *et al.* Global analysis of protein activities using proteome chips. *Science* **293**, 2101–2105 (2001).
- Zhu, H. *et al.* Analysis of yeast protein kinases using protein chips. *Nature Genet.* **26**, 283–289 (2000).
- Gavin, A. C. *et al.* Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
- Ito, T. *et al.* A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl Acad. Sci. USA* **98**, 4569–4574 (2001).
- Uetz, P. *et al.* A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623–627 (2000).
- Xenarios, I. *et al.* DIP: the database of interacting proteins. *Nucleic Acids Res.* **28**, 289–291 (2000).
- Bader, G. D. & Hogue, C. W. BIND—a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. *Bioinformatics* **16**, 465–477 (2000).
- Horak, C. E. *et al.* Complex transcriptional circuitry at the G1/S transition in *Saccharomyces cerevisiae*. *Genes Dev.* **16**, 3017–3033 (2002).
- Lee, T. I. *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
- Toda, T., Cameron, S., Sass, P., Zoller, M. & Wigler, M. Three different genes in *S. cerevisiae* encode the catalytic subunits of the cAMP-dependent protein kinase. *Cell* **50**, 277–287 (1987).
- Pan, X. & Heitman, J. Cyclic AMP-dependent protein kinase regulates pseudohyphal differentiation in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 4874–4887 (1999).
- Jonassen, I., Collins, J. F. & Higgins, D. G. Finding flexible patterns in unaligned protein sequences. *Protein Sci.* **4**, 1587–1595 (1995).
- Hanrahan, J. & Snyder, M. Cytoskeletal activation of a checkpoint kinase. *Mol. Cell* **12**, 663–673 (2003).
- Barral, Y., Parra, M., Bidlingmaier, S. & Snyder, M. Nim1-related kinases coordinate cell cycle progression with the organization of the peripheral cytoskeleton in yeast. *Genes Dev.* **13**, 176–187 (1999).
- Mewes, H. W., Albermann, K., Heumann, K., Liebl, S. & Pfeiffer, F. MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acids Res.* **25**, 28–30 (1997).
- Ubersax, J. A. *et al.* Targets of the cyclin-dependent kinase Cdk1. *Nature* **425**, 859–864 (2003).
- Loog, M. & Morgan, D. O. Cyclin specificity in the phosphorylation of cyclin-dependent kinase substrates. *Nature* **434**, 104–108 (2005).

21. Gruhler, A. *et al.* Quantitative phosphoproteomics applied to the yeast pheromone signalling pathway. *Mol. Cell Proteomics* **4**, 310–327 (2005).
22. Huse, M. & Kuriyan, J. The conformational plasticity of protein kinases. *Cell* **109**, 275–282 (2002).
23. Moffat, J. & Andrews, B. Late-G1 cyclin-CDK activity is essential for control of cell morphogenesis in budding yeast. *Nature Cell Biol.* **6**, 59–66 (2004).
24. Breitkreutz, B. J., Stark, C. & Tyers, M. Osprey: a network visualization system. *Genome Biol.* **4**, R22.1–4 (2003).
25. Huh, W. K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
26. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
27. Winzeler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Tang for help with the initial Cdc28 preparations, and D. Gelperin, J. Mok and K. Wise for comments on the manuscript. M.S. was funded by grants from the NIH; J.P., G.D. and J.F. were

funded by NIH predoctoral fellowships, and B.A. and M.T. were funded by grants from the Canadian Institutes of Health Research. M.J.R.S. was funded by a project grant from the Wellcome Trust, UK.

Author Contributions Assay development was performed by H.Z., J.P., G.D., G.M. and M.S. Proteome chips were prepared by G.M., B.S. and P.F.P. at Invitrogen. G.J. contributed transcription factors for the arrays. Kinase assays were performed by J.P., G.D., H.Z. and M.S. Most kinases were prepared by J.P. and G.D. Additional kinases were provided by A.B., R.S., R.R.M., M.C.S., N.R., S.J.L., A.S.M., M.J.R.S., D.F.S., C.D.V., M.T. and B.A. Data analysis was performed by X.Z., J.P. and G.D. Consensus mapping was by H.G. *In vitro* solution validations were performed by G.M. and L.M. *In vivo* substrate validations were performed by J.P., G.D. and J.F. Most assays and analyses were performed in the laboratory of M.S. with contributions from M.G.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper at www.nature.com. Correspondence and requests for materials should be addressed to M.S. (michael.snyder@yale.edu).

Silencing of microRNAs *in vivo* with 'antagomirs'

Jan Krützfeldt¹, Nikolaus Rajewsky³, Ravi Braich⁴, Kallanthottathil G. Rajeev⁴, Thomas Tuschl², Muthiah Manoharan⁴ & Markus Stoffel¹

MicroRNAs (miRNAs) are an abundant class of non-coding RNAs that are believed to be important in many biological processes through regulation of gene expression^{1–3}. The precise molecular function of miRNAs in mammals is largely unknown and a better understanding will require loss-of-function studies *in vivo*. Here we show that a novel class of chemically engineered oligonucleotides, termed 'antagomirs', are efficient and specific silencers of endogenous miRNAs in mice. Intravenous administration of antagomirs against miR-16, miR-122, miR-192 and miR-194 resulted in a marked reduction of corresponding miRNA levels in liver, lung, kidney, heart, intestine, fat, skin, bone marrow, muscle, ovaries and adrenals. The silencing of endogenous miRNAs by this novel method is specific, efficient and long-lasting. The biological significance of silencing miRNAs with the use of antagomirs was studied for miR-122, an abundant liver-specific miRNA. Gene expression and bioinformatic analysis of messenger RNA from antagomir-treated animals revealed that the 3' untranslated regions of upregulated genes are strongly enriched in miR-122 recognition motifs, whereas downregulated genes are depleted in these motifs. Analysis of the functional annotation of downregulated genes specifically predicted that cholesterol biosynthesis genes would be affected by miR-122, and plasma cholesterol measurements showed reduced levels in antagomir-122-treated mice. Our findings show that antagomirs are powerful tools to silence specific miRNAs *in vivo* and may represent a therapeutic strategy for silencing miRNAs in disease.

Approaches to the study of miRNA function in mammals have focused on the overexpression or inhibition of miRNAs with antisense 2'-O-methyl (2'-OMe) oligoribonucleotides in cell lines as well as computational target predictions and validation by luciferase reporter assays^{4–8}. However, functional studies in mice that lack specific miRNAs have yet to be reported. Further, because miRNAs may be important in human disease^{6,9–12}, approaches to interrupt miRNA function may have therapeutic utility. Small interfering double-stranded RNAs (siRNAs) engineered with certain 'drug-like' properties such as chemical modifications for stability and cholesterol conjugation for delivery have been shown to achieve therapeutic silencing of an endogenous gene *in vivo*¹³. To develop a pharmacological approach for silencing miRNAs *in vivo*, we designed chemically modified, cholesterol-conjugated single-stranded RNA analogues complementary to miRNAs, and have termed these oligonucleotides 'antagomirs'.

To explore the potential of these synthetic RNA analogues to silence endogenous miRNAs, we designed an antagomir, antagomir-122, selective for miR-122, an abundant, liver-specific miRNA. Antagomir-122 was synthesized starting from a hydroxyprolinol-linked cholesterol solid support¹⁴ and 2'-OMe phosphoramidites. This compound was administered to mice by intravenous injection in a small volume (0.2 ml) at normal pressure. Administration of

antagomir-122 resulted in a marked decrease in endogenous miR-122 levels as detected by northern blots (Fig. 1a). Administration of unmodified single-stranded RNA (anti-122) had no effect on levels of hepatic miR-122 (Fig. 1a), whereas injection of unconjugated, but chemically modified, single-stranded RNAs with a

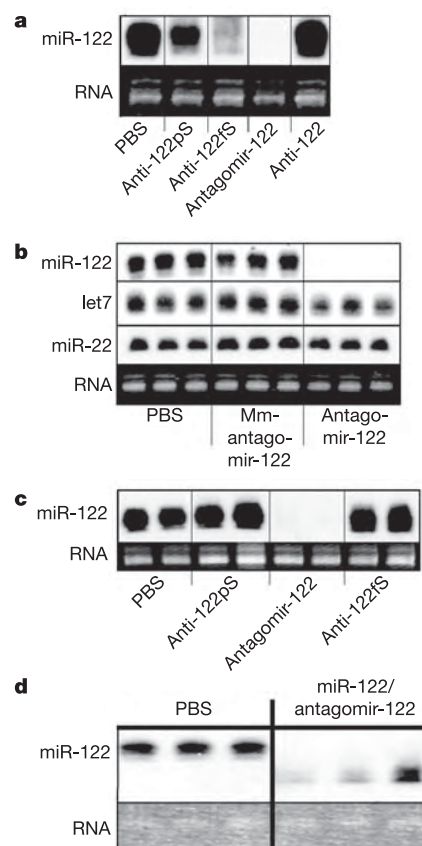


Figure 1 | Specific targeting of miR-122 in mouse liver by tail-vein injection of chemically modified single-stranded RNAs. **a**, Northern blots of total RNA (15 µg) isolated from mouse liver 24 h after injection of differently modified RNAs (three injections of 80 mg kg⁻¹ d⁻¹) against miR-122. Samples were separated in 14% polyacrylamide gels in the absence (**a**, **b**) or presence (**c**) of 20% formamide. Ethidium bromide staining of tRNA is shown as a loading control. **d**, Mice were injected intravenously with PBS or a miR-122–antagomir-122 duplex (three injections of 80 mg kg⁻¹ d⁻¹) and the livers were harvested 24 h after the last injection. Expression of miRNA-122 was analysed by northern blotting.

¹Laboratory of Metabolic Diseases, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA. ²Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA. ³Biology and Mathematics, Center for Comparative Functional Genomics, Department of Biology, New York University, New York, New York 10003, USA. ⁴Anylam Pharmaceuticals Inc., 300 3rd Street, Cambridge, Massachusetts 02142, USA.

partially modified (pS) or fully modified (fS) phosphorothioate backbone and 2'-OMe sugar modifications (anti-122pS and anti-122fS, respectively) led to an incomplete effect (Fig. 1a). The effects of antagomir-122 were found to be specific because animals injected with a control for antagomir-122 that harboured four mismatch mutations (mm-antagomir-122) had no effect on miR-122 levels (Fig. 1b). Furthermore, expression levels of let7 and miR-22 were unaffected in mice treated with antagomir-122 and mm-antagomir-122, indicating that silencing was miRNA-specific (Fig. 1b).

MicroRNA-122 is expressed at high levels in hepatocytes, with more than 50,000 copies per cell¹⁵. To examine whether the decrease in miR-122 after treatment with antagomir was due to degradation or to stoichiometric duplex formation between miR-122 and antagomir-122, with the resulting inability of the probe to detect miR-122 in a northern blot, we analysed total RNA from livers of mice treated with unconjugated anti-miR-122 oligonucleotides (anti-122fS and anti-122pS) or antagomir-122 under stringent denaturing conditions (Fig. 1c). No decrease in miR-122 levels could be detected with unconjugated anti-miR-122 RNA-treated livers, showing that the decrease in miR-122 levels observed under non-denaturing conditions (Fig. 1a) was simply due to the formation of miR-122–RNA duplexes. In contrast, miR-122 remained undetectable in livers of mice treated with antagomir-122 and was not lost during the RNA extraction procedure (Supplementary Figs 1 and 2). Further, we were able to detect specific miR-122 degradation products when the amount of miR-122 was increased by delivering preformed miR-122–antagomir-122 duplexes into the liver (Fig. 1d). These data indicate that the silencing of miRNA-122 in mice treated with antagomir-122 might have been due to degradation of the miRNA. The ability of antagomir-122 treatment to result in miR-122 degradation is probably due to the efficient delivery to hepatocytes and/or a consequence of changes in subcellular localization of antagomir–miR-122 complexes.

Several pharmacological properties for antagomirs were evaluated further, including dose–response, duration of action, and biodistribution. To determine the dose of antagomir-122 that can completely silence miR-122 we injected mice with a total dose of 80, 160 or 240 mg per kg body weight and analysed for miR-122 expression levels. The highest dose resulted in a complete loss of miR-122 signal (Supplementary Fig. 3a). We also tested the duration of silencing that could be achieved after the injection of antagomir-122. Levels of miR-122 were undetectable for as long as 23 days after injection, indicating that silencing of miRNAs using antagomirs is long lasting (Supplementary Fig. 3b). We next investigated the bioavailability and silencing activity of antagomirs in different tissues, using miR-16 as a target because it is expressed abundantly in all tissues. In mice treated with antagomir-16, miR-16 was efficiently silenced in all tissues tested except brain (Fig. 2a). Antagomir-16 did not seem to affect the expression of the 89-nucleotide precursor of miR-16 detected in bone marrow. The bioavailability of antagomir-16 was also assessed by northern blotting in the above-mentioned tissue samples. Consistent with the ability to silence miR-16 was our detection of significant levels of antagomir-16 in all tissues except brain (Fig. 2b). These data show that antagomirs achieve broad bio-distribution and can efficiently silence miRNAs in most tissues *in vivo*.

Many miRNA genes have been found to be located close together and to be coordinately transcribed. These polycistronic miRNA genes result in long primary transcripts (pri-miRNAs) that are processed by multiple enzymes in the nucleus and cytoplasm to generate the mature miRNAs¹⁶. To investigate whether antagomirs targeting polycistronic miRNAs retain their target specificity with no effect on the expression of co-transcribed miRNAs, we injected mice with antagomirs targeting either miR-192 or miR-194 of the bicistronic cluster miR-192/194. Administration of antagomir-192 into mice resulted in the silencing of miR-192 in liver and kidney, with no effect on the expression levels of miR-194. The converse effects were seen

with injection of antagomir-194 (Supplementary Fig. 4). These data show that antagomirs have the ability to differentially silence specific miRNAs that derive from the same primary transcript.

MicroRNAs can regulate the mRNA levels of their targets^{17,18}, and pharmacological silencing of miRNAs using antagomirs might therefore lead to the regulation of many mRNAs. To identify genes that are regulated by miR-122 we performed gene-expression analysis with Affymetrix arrays in livers from mice treated with antagomir-122 and mm-antagomir-122. In all, 363 transcripts were upregulated (at least 1.4-fold) and 305 transcripts were downregulated (at least 1.4-fold) in antagomir-122-treated mice compared with controls (Supplementary Table 1). The regulation of genes that were upregulated was confirmed by reverse-transcriptase-mediated polymerase chain reaction (RT–PCR; Fig. 3a). These included those members of gene families that are usually repressed in hepatocytes, including those

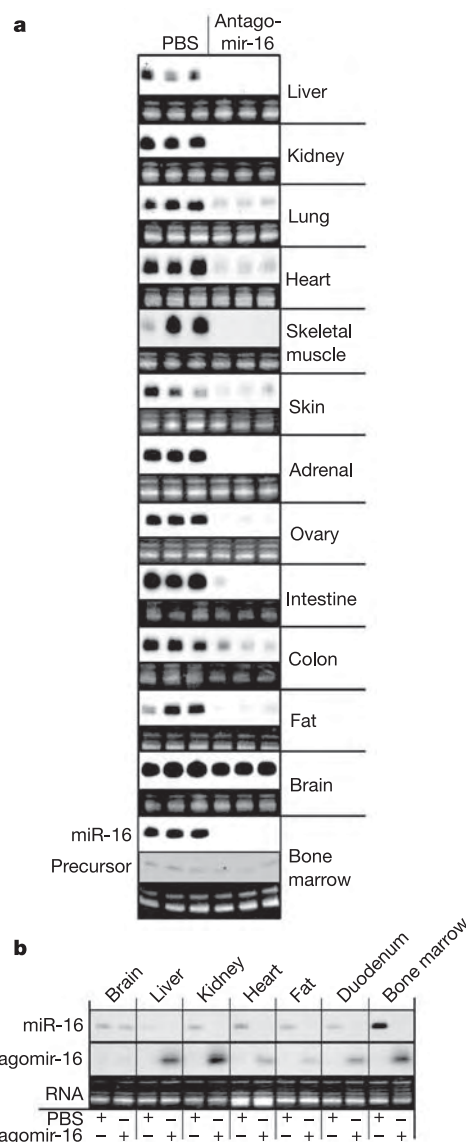


Figure 2 | Antagomirs target microRNA expression in multiple tissues. **a**, Northern blots of total RNA (10–30 µg) isolated from different mouse tissues 24 h after injection of antagomir-16 ($n = 3$). The precursor miRNA was visible on northern blots of bone marrow. Membranes were probed for miR-16. **b**, Total RNA from three mice as shown in **a** were pooled for the detection of miR-16 and the injected antagomir-16. Ethidium bromide staining of tRNA is shown as a loading control.

encoding aldolase-A (aldo-A), N-Myc downstream regulated gene 3 (Ndr3) and IQ-motif-containing GTPase-activating protein 1 (Iqgap1). Therefore, miR-122 could contribute to the maintenance of the adult liver phenotype, as previously suggested for two other tissue-specific miRNAs¹⁷. To assess the motif contents of significantly upregulated and downregulated genes further, we analysed the 3' untranslated region (UTR) sequences of 9,554 mRNAs that have annotated 3' UTRs. Of these, 142 mRNAs were statistically significantly upregulated with at least a 1.4-fold change. Figure 3b shows the percentage of genes that had at least one miR-122 recognition motif CACTCC, corresponding to nucleotides 2–7 of miR-122—the core 'nucleus' sequence¹⁹ (also referred to as the 'seed' sequence²⁰). We observed a highly significant 2.6-fold increase in the probability of having at least one miR-122 nucleus in the 3' UTR of upregulated

transcripts in comparison with genes with no change in mRNA levels (Fig. 3b). Many of the miR-122 nucleus sequences in upregulated genes had not previously been predicted^{19–21}, indicating that the number of direct miRNA targets might be significantly larger than previously estimated.

For experimental validation of the link between repression and the presence of miR-122 nucleus matches within the 3' UTR, we cloned the 3' UTR of five genes upregulated by antagomir-122 and containing a miR-122 nucleus sequence into a luciferase reporter system. When co-transfected with miR-122, all reporters exhibited significant repression relative to co-transfections with control miRNA, showing that miR-122 binding to its nucleus contributes directly to mRNA repression (Fig. 3c). Of 108 transcripts that were significantly downregulated, we observed that the probability of harbouring a

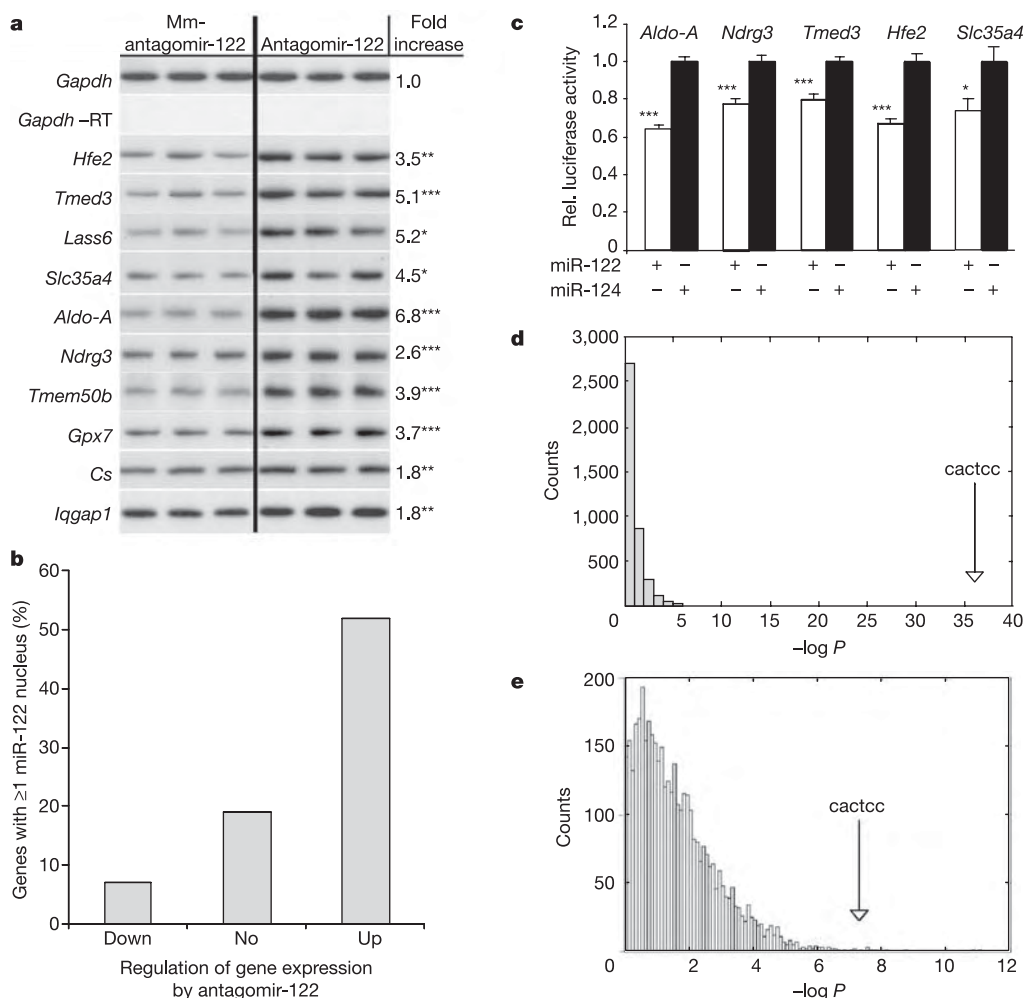


Figure 3 | Positive and negative regulation of gene expression by miRNA-122. **a**, Steady-state mRNA levels of genes in livers of mice treated with mm-antagomir-122 or antagomir-122. Expression was measured by RT-PCR. Each lane indicates an individual animal. Gene symbols are shown in accordance with the International Standardized Nomenclature (www.informatics.jax.org/mgihome/nomen/gene.shtml). Fold increase indicates the ratio of expression levels of the means of mice treated with antagomir-122 and mm-antagomir-122. The glyceraldehyde-3-phosphate dehydrogenase gene (*Gapdh*) was used as a loading control. **b**, Abundance of the miR-122 nucleus in differentially expressed genes. The plot shows the percentage of genes with at least one miR-122 recognition motif present in their 3' UTR. The significance of the lower (higher) percentage of downregulated (upregulated) genes with a miR-122 nucleus was assessed by 1,000 random poolings of the same number of genes for each class from the total set of transcripts. The result for downregulated genes was significant at

three standard deviations ($P = 0.001$). The result for upregulated genes was significant at more than eight standard deviations. **c**, Micro-RNA-directed repression of *Renilla* luciferase reporter genes bearing 3' UTR segments from genes identified from microarray expression analysis of antagomir-122-treated mice after co-transfection into HEK-293 cells with the indicated miRNA (si-122 or si-124). Data are from three independent experiments and are shown as means $\pm$ s.e.m., with $n = 6$. **d**, For each of the 4,096 possible hexamer RNA motifs and each transcript, the number of non-overlapping occurrences divided by the length of the 3' UTR was recorded. For each motif, a non-parametric test (the one-tailed Wilcoxon rank sum test) was applied to these distributions in upregulated versus no-change transcripts. Shown is the histogram of the negative natural logarithm of all 4,096 P values. **e**, Analogously to **d**, comparison between transcripts with no significant change in expression and downregulated transcripts. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks $P < 0.001$.

miR-122 nucleus was decreased by almost the same factor, 2.7-fold (Fig. 3b). To analyse further whether the over-representation and under-representation of miR-122 nucleus sequences are specific, we analysed the abundance of all 4,096 possible hexamer motifs across downregulated, upregulated and unchanged transcripts. When we compared upregulated with unchanged genes, the miR-122 nucleus sequence was the most significantly over-represented hexamer

(Fig. 3d). Notably, the miR-122 nucleus was within the top 1% of under-represented motifs for downregulated transcripts (Fig. 3e), indicating a possible evolutionary tendency of downregulated genes to lack binding sites for miR-122. These results indicate that upregulated mRNAs are directly targeted and repressed by miR-122, but also that a significant number of downregulated genes are likely to be activated by miR-122.

To assess the phenotype of altered gene regulation by miR-122, we analysed the annotation of regulated genes for enrichment in Gene Ontology categories. The top-ranking functional category was 'cholesterol biosynthesis' ($P = 1.6 \times 10^{-11}$) and was found for gene transcripts downregulated by antagomir-122. The expression of at least 11 genes involved in cholesterol biosynthesis was decreased between 1.4-fold and 2.3-fold in antagomir-122-treated mice (Supplementary Table 1); some of these were confirmed by RT-PCR (Fig. 4a). Mice injected with an adenovirus expressing miR-122 (Ad-122) increased the expression of some of these genes (Fig. 4b). One of the gene transcripts downregulated by treatment with antagomir-122 was 3-hydroxy-3-methylglutaryl-CoA-reductase (Hmgcr), a rate-limiting enzyme of endogenous cholesterol biosynthesis. We measured the enzymatic activity of Hmgcr in liver extracts and found a roughly 45% decrease in activity in antagomir-122-treated mice compared with mm-antagomir-122-treated mice (9.7 ± 1.0 and 17.2 ± 2.3 pmol per mg of microsomal protein per minute, respectively; $n = 4$, $P = 0.02$). Consistent with this effect was the observation that plasma cholesterol levels were decreased more than 40% in treated animals whereas there was no detectable effect on plasma non-esterified fatty acids, triglyceride, bile acid and glucose levels (Fig. 4c). No decrease in cholesterol was observed with antagomir-16, antagomir-192 and antagomir-194, showing that the effects of antagomir-122 are sequence-specific and unrelated to the use of a cholesterol-conjugated oligonucleotide itself. Decreased cholesterol levels in antagomir-122-treated mice lasted for at least 2 weeks (data not shown). Antagomir-122 was well tolerated even during the course of the prolonged treatment; no alterations in body weight or serum markers of liver toxicity (alanine aminotransferase (ALT) levels) were detected. Together, these data indicate that miR-122 participates in regulation of the cholesterol biosynthetic pathway and that silencing of a miRNA can be achieved without apparent toxicities.

The discovery of miRNAs is likely to change our understanding of gene expression fundamentally, yet almost nothing is known of their function in mammalian systems *in vivo*. Our data show that antagomirs can effectively silence miRNAs *in vivo*, and that antagomirs can enable the study of gene regulation *in vivo* by a tissue-specific miRNA, miR-122. In addition to many upregulated genes, silencing of miR-122 also led to a decrease in a significant number of genes. The mechanism by which miRNAs can activate gene expression is currently unknown but may be the result of a direct effect (chromatin remodelling) or an indirect effect (suppression of a transcriptional repressor). Notwithstanding the large number of genes regulated by antagomir treatment, it is striking, given the extent and duration of miR-122 silencing, that the phenotype of antagomir-122-treated mice was otherwise modest. Further studies are needed to explore changes at the protein level as a result of silencing multiple miRNAs as well as potential effects under stress conditions or in disease models, but our data support the model that some miRNAs might serve more as a 'rheostat' than as an 'on-off switch', to fine-tune gene expression. Our novel pharmacological approach to silence miRNAs specifically will allow the rapid generation of mice lacking specific miRNAs or combinations of miRNAs for further functional studies. Finally, because it has been shown that miRNAs are involved in cancer^{9–12}, cell growth and differentiation^{4,5,22}, insulin secretion⁶ and viral infection²³, silencing of miRNAs with antagomirs could become a therapeutic strategy for diseases²⁴ such as cancer, hepatitis and diabetes, and others almost certain to be discovered, in which miRNAs participate in disease aetiology.

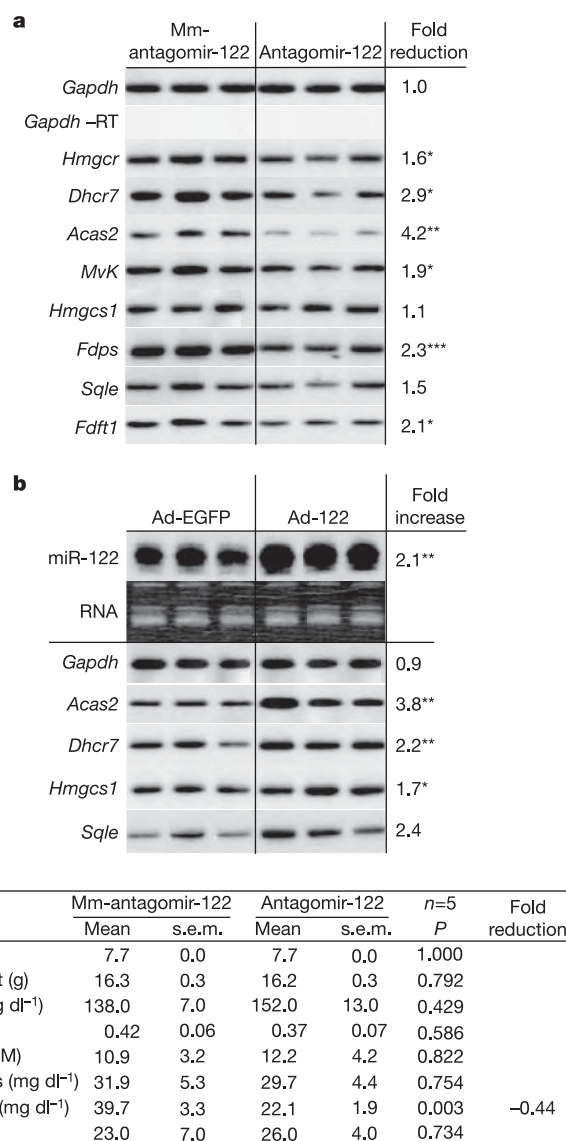


Figure 4 | MicroRNA-122 regulates the expression of genes involved in cholesterol biosynthesis. **a**, RT-PCR analysis of cholesterol biosynthesis genes identified in Affymetrix gene-expression analysis in livers of mice treated with antagomir-122 or mm-antagomir-122. Fold reduction indicates the ratio of expression levels of the means of mice treated with mm-antagomir-122 and antagomir-122. The glyceraldehyde-3-phosphate dehydrogenase gene (*Gapdh*) was used as a loading control. Gene abbreviations: *Hmgcr*, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; *Dhcr7*, 7-dehydrocholesterol reductase; *Acas2*, acetyl-coenzyme A synthetase 2 (ADP forming); *Mvk*, mevalonate kinase; *Hmgcs1*, 3-hydroxy-3-methylglutaryl-coenzyme A synthase 1; *Fdps*, farnesyl diphosphate synthetase; *Sqle*, squalene epoxidase; *Fdft1*, farnesyl diphosphate farnesyl transferase 1. Each lane represents an individual animal. **b**, RT-PCR analysis of cholesterol biosynthesis genes in animals 6 days after injection of Ad-EGFP or Ad-122. The upper row shows a northern blot of liver RNA for miR-122. **c**, Metabolic measurements in mice treated with antagomir-122 and mm-antagomir-122 control. FFA, non-esterified fatty acid. Asterisk, $P < 0.05$; two asterisks, $P < 0.01$; three asterisks, $P < 0.001$.

METHODS

Synthesis of antagomirs. The single-stranded RNAs and modified RNA analogues used in this study consisted of a 21–23-nucleotide length with modifications as specified: anti-122, 5′-ACAAACACCAUUGUCACACUCCA-3′; anti-122pS, 5′-a₃c₃aaacacauugucacac₃u₃c₃a-3′; anti-122fS, 5′-a₃c₃a₃a₃a₃c₃a₃u₃u₃g₃u₃c₃a₃c₃a₃c₃u₃c₃a-3′; antagomir-122, 5′-a₃c₃aaacacacauugucacac₃c₃a₃-Chol-3′; mm-antagomir-122, 5′-a₃c₃acacacacugucacauu₃c₃a₃-Chol-3′; antagomir-16, 5′-c₃g₃ccaaauuuuacuguc₃c₃u₃a₃-Chol-3′; antagomir-192, 5′-g₃g₃cugucacauuacaggu₃c₃a₃g₃-Chol-3′; antagomir-194, 5′-u₃c₃cacacaggaugucuu₃a₃c₃a₃-Chol-3′. The lower case letters represent 2′-OMe-modified nucleotides; subscript 's' represents a phosphorothioate linkage; 'Chol' represents cholesterol linked through a hydroxypropylolinkage¹⁴; anti-122pS is anti-miR-122 RNA with partial phosphorothioate backbone and anti-122fS is anti-miR-122 RNA with full phosphorothioate backbone. Details of the synthesis are given in the Supplementary Methods.

Animals. All animal models were maintained in a C57BL/6J background on a 12-h light/dark cycle in a pathogen-free animal facility at Rockefeller University. Six-week-old mice received, on one to three consecutive days, tail-vein injections of saline or different RNAs (as indicated). RNAs were administered at doses of 80 mg per kg body weight in 0.2 ml per injection. Measurements of miRNA levels in tissues were performed 24 h after the last injection unless indicated otherwise. Tissues were harvested, snap-frozen and stored at −80 °C.

Generation of recombinant adenovirus. The recombinant adenovirus used to express miR-122 (Ad-122) was generated by PCR, amplifying a 344-base-pair miRNA precursor sequence with primers 5′-AGTCAGATGTACAGTTATAAGCACAGAGGACCAG-3′ and 5′-TTATTCAAGATCCCGGGGCTCTTCC-3′. The fragment was cloned into vector Ad5CMV-KnpA. Ad-EGFP (ViraQuest) was used as a control. Mice were infected with 5 × 10⁹ plaque-forming units per mouse by injection into the tail vein.

Gene-expression analysis. Total RNA from liver of mice treated with antagomirs was isolated three days after the initiation of treatment. RNA was pooled from three animals for each group. The generation and analysis of Affymetrix microarray data are described in the Supplementary Methods.

Northern blotting analysis. Total RNA was isolated with the Trizol reagent (Invitrogen) and precipitation with ethanol. RNA was separated at 45 mA on 14% polyacrylamide gels containing 8 M urea and 20% formamide in a Protean II xi vertical electrophoresis cell (Bio-Rad). Samples were transferred to Hybond-N⁺ nylon membranes (Amersham) in a Trans-Blot electrophoretic transfer cell (Bio-Rad) for 2 h at 1 A. DNA antisense probes were designed as described in the 'microRNA registry' (<http://www.sanger.ac.uk>) and 20 pmol of each was labelled with T4 polynucleotide kinase (New England Biolabs) and 30 μCi of [γ-³²P]ATP (3,000 Ci mmol^{−1}; NEN Life Science). Hybridizations were performed at 50 °C for 16 h in a 20 mM sodium phosphate buffer pH 7.2 containing 7% SDS, 0.75 M NaCl, 75 mM sodium citrate, 0.02% albumin, 0.02% polyvinylpyrrolidone, 0.02% Ficoll 400 and 0.1 mg ml^{−1} sonicated salmon-sperm DNA.

RT-PCR. Extraction of total RNA, synthesis of cDNA, and PCR were performed as described in the Supplementary Methods.

Assay of luciferase activity. Mouse full-length 3′ UTR sequences were PCR-amplified with specific primers and cloned downstream of the stop codon in pRL-TK (Promega). HEK-293 cells were cultured in 24-well plates and each well was transfected with 50 ng of the respective pRL-TK 3′ UTR constructs (Rr-luc), 50 ng of pGL3 control vector (Pp-luc) (Promega) and 200 ng of double-stranded siRNA (Dharmacon). Cells were harvested and assayed 24–30 h after transfection. Results were normalized to the Pp-luc control and are expressed relative to the average value of the control miRNA (si-124).

3′ UTR sequences and mapping of array probes to transcripts. We extracted mouse 3′ UTRs with the use of the Refseq data set (<ftp://ftp.ncbi.nih.gov/refseq/>)²⁵. A total of 17,264 3′ UTR sequences at least 30 nucleotides in length were obtained. Affymetrix probe identifiers were assigned to the Refseq transcripts by using a mapping provided by Ensembl (<http://www.ensembl.org/Multi/martview>). When more than one probe identifier mapped to a transcript, we insisted that the Affymetrix significance call be consistent for all probes. Transcripts were discarded otherwise. The fold change assigned to a transcript was the average of all probes that mapped to the transcript. Finally, a cut-off of 0.5 in the log₂ of fold changes was applied.

Gene ontology analysis. The analysis is described in the Supplementary Methods.

Hmgcr activity assay. Hepatic microsomal Hmgcr activity was assayed by a method modified from a previously published procedure and described in detail in the Supplementary Methods.

Statistical analysis. Results are given as means ± s.e.m. Statistical analyses were

performed with Student's *t*-test, and the null hypothesis was rejected at the 0.05 level.

Received 19 July; accepted 12 October 2005.

Published online 30 October 2005.

- Ambros, V. The functions of animal microRNAs. *Nature* **431**, 350–355 (2004).
- Bartel, D. P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- He, L. & Hannon, G. J. MicroRNAs: Small RNAs with a big role in gene regulation. *Nature Rev. Genet.* **5**, 522–531 (2004).
- Zhao, Y., Samal, E. & Srivastava, D. Serum response factor regulates a muscle-specific microRNA that targets Hand2 during cardiogenesis. *Nature* **436**, 214–220 (2005).
- Chen, C. Z. & Lodish, H. F. MicroRNAs as regulators of mammalian hematopoiesis. *Semin. Immunol.* **17**, 155–165 (2005).
- Poy, M. N. *et al.* Pancreatic islet-specific microRNA regulates insulin secretion. *Nature* **432**, 226–230 (2004).
- Meister, G., Landthaler, M., Dorsett, Y. & Tuschl, T. Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. *RNA* **10**, 544–550 (2004).
- Hutvagner, G., Simard, M. J., Mello, C. C. & Zamore, P. D. Sequence-specific inhibition of small RNA function. *PLoS Biol.* **2**, E98 (2004).
- Eis, P. S. *et al.* Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc. Natl. Acad. Sci. USA* **102**, 3627–3632 (2005).
- Metzler, M., Wilda, M., Busch, K., Viehmann, S. & Borkhardt, A. High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. *Genes Chromosom. Cancer* **39**, 167–169 (2004).
- He, L. *et al.* A microRNA polycistron as a potential human oncogene. *Nature* **435**, 828–833 (2005).
- Lu, J. *et al.* MicroRNA expression profiles classify human cancers. *Nature* **435**, 834–838 (2005).
- Soutschek, J. *et al.* Therapeutic silencing of an endogenous gene by systemic administration of modified siRNAs. *Nature* **432**, 173–178 (2004).
- Manoharan, M., Kesavan, V., & Rajeev, K. G. siRNA's containing ribose substitutes to which lipophilic moieties may be attached. U.S. Pat. Appl. Publ. (2005), US 2005107325.
- Chang, J. *et al.* miRNA-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT1. *RNA Biol.* **1**, 106–113 (2004).
- Tomari, Y. & Zamore, P. D. MicroRNA biogenesis: drosha can't cut it without a partner. *Curr. Biol.* **15**, R61–R64 (2005).
- Lim, L. P. *et al.* Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* **433**, 769–773 (2005).
- Bagga, S. *et al.* Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell* **122**, 553–563 (2005).
- Krek, A. *et al.* Combinatorial microRNA target prediction. *Nature Genet.* **37**, 495–500 (2005).
- Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
- John, B. *et al.* Human MicroRNA targets. *PLoS Biol.* **2**, e363 (2004).
- Cheng, A. M., Byrom, M. W., Shelton, J. & Ford, L. P. Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis. *Nucleic Acids Res.* **33**, 1290–1297 (2005).
- Jopling, C. L., Yi, M., Lancaster, A. M., Lemon, S. M. & Sarnow, P. Modulation of hepatitis C virus RNA abundance by a liver-specific microRNA. *Science* **309**, 1577–1580 (2005).
- Weiler, J., Hunziker, J. & Hall, J. Anti-miRNA oligonucleotides (AMOs): ammunition to target miRNAs implicated in human disease? *Gene Therapy* Advance online publication 29 September 2005; doi:10.1038/sj.gt.3302654.
- Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Maraganore, V. Kotelianski and P. Sharp for discussion and suggestions. These studies were supported by NIH grants (to M.S., T.T. and N.R.), and an unrestricted grant from Bristol Myers Squibb (M.S.).

Author Information Gene expression data have been deposited at GEO-NCBI under accession number GSE3425. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.S. (stoffel@rockefeller.edu).

LETTERS

The APC/C and CBP/p300 cooperate to regulate transcription and cell-cycle progression

Andrew S. Turnell¹, Grant S. Stewart^{1,2}, Roger J. A. Grand¹, Susan M. Rookes¹, Ashley Martin¹, Hiroyuki Yamano³, Stephen J. Elledge^{2,4} & Phillip H. Gallimore¹

The anaphase-promoting complex/cyclosome (APC/C) is a multi-component E3 ubiquitin ligase that, by targeting protein substrates for 26S proteasome-mediated degradation through ubiquitination, coordinates the temporal progression of eukaryotic cells through mitosis and the subsequent G1 phase of the cell cycle^{1–4}. Other functions of the APC/C are, however, less well defined. Here we show that two APC/C components, APC5 and APC7, interact directly with the coactivators CBP and p300 through protein–protein interaction domains that are evolutionarily conserved in adenovirus E1A^{5–8}. This interaction stimulates intrinsic CBP/p300 acetyltransferase activity and potentiates CBP/p300-dependent transcription. We also show that APC5 and APC7 suppress E1A-mediated transformation in a CBP/p300-dependent manner, indicating that these components of the APC/C may be targeted during cellular transformation. Furthermore, we establish that CBP is required in APC/C function; specifically, gene ablation of CBP by RNA-mediated interference markedly reduces the E3 ubiquitin ligase activity of the APC/C and the progression of cells through mitosis. Taken together, our results define discrete roles for the APC/C–CBP/p300 complexes in growth regulation.

Characterization of the APC/C has shown that it is a cullin-based macromolecular complex that uses the ring-finger component APC11 as its E3 ubiquitin ligase^{1,2}. Sequence analyses have determined that many APC/C subunits possess TPR protein–protein interaction motifs that mediate the cell-cycle-dependent association of the APC/C with the activator proteins Cdc20 and Cdh1 (ref. 9). In studies on the relationship between E1A and the ubiquitin–proteasome pathway^{10,11}, we determined that the APC/C components APC5 and APC7 have considerable sequence homology with the amino-terminal region of E1A (Fig. 1a, b) and a section of E1A comprising part of conserved region 1 (CR1; Fig. 1c). As E1A function is highly modular, targeting cellular binding partners through short highly conserved motifs⁸, these data suggested that E1A, APC5 and APC7 might target common cellular proteins. Indeed, reciprocal immunoprecipitation studies showed that E1A-binding proteins CBP and p300 associated *in vivo* with both APC5 and APC7 (Fig. 1d). Crucially, antisera against APC2 and APC6 similarly coimmunoprecipitated CBP/p300, suggesting that the APC/C holoenzyme forms stable complexes with CBP and p300 *in vivo* (Fig. 1d, western blot). CBP/p300 association with the APC/C is not, however, cell-cycle regulated (Supplementary Fig. S1).

To establish whether APC5 and APC7 bind CBP/p300 independently, pull-downs assays were done with glutathione S-transferase (GST) fusion proteins of CBP and *in vitro* translated, [³⁵S]L- α -methionine labelled APC5 and APC7. We found that APC5 and APC7 bound independently to CBP, predominantly through C/H3, a

region that is also targeted by E1A⁵ (Fig. 1e). APC5 and APC7 similarly bound to the N-terminal region of CBP, albeit with lower affinity than E1A. In contrast to E1A, however, APC5 and APC7 also bound to the acetyltransferase domain of CBP (Fig. 1e). Consistent with the notion that the E1A homology domains in APC5, APC7 and E1A define principal CBP/p300-binding sites, mutants lacking either the N-terminal or the CR1 homology domain had a considerably lower affinity for C/H3 than did full-length proteins (Supplementary Fig. S2a, b).

The requirement for CBP/p300 in transcriptional activation^{5,12–21} led us to investigate whether APC5 and/or APC7 affected CBP/p300-dependent transcription. We made use of the fact that when tethered to the DNA-binding domain of Gal4, CBP/p300 will stimulate transcription from Gal4-responsive promoters in the absence of the Gal4 activation domain¹⁵ (Fig. 2a). Intriguingly, both APC5 and APC7 potentiated p300-dependent transactivation, whereas E1A expression inhibited p300-dependent transactivation (Fig. 2a). APC5 or APC7 mutants, lacking either the N-terminal or CR1 homology domains were substantially impaired in their ability to potentiate transactivation (Fig. 2a). Consistent with these findings, E1A mutants lacking the N terminus or mutated in CR1 did not inhibit p300-dependent transactivation (Fig. 2a).

Because CBP/p300 can modulate p53 transcriptional activity^{18–20}, we investigated whether APC5 and/or APC7 similarly affected p53 activity. Notably, both APC5 and APC7 enhanced p53-dependent transactivation of a p21^{CIP1/WAF1} promoter-tethered reporter²¹ in a p300-dependent manner (Fig. 2b). To determine whether APC5 and/or APC7 regulate CBP/p300-dependent transcription *in vivo*, we investigated whether ablation of APC5 or APC7 gene expression by RNA-mediated interference (RNAi) affected the p53-dependent transactivation of the p21^{CIP1/WAF1} promoter in primary human skin fibroblasts (HSFs) after treatment with ionizing radiation. Basal levels of p21^{CIP1/WAF1} messenger RNA were reduced after APC5 and/or APC7 knockdown relative to non-silencing controls (Fig. 2c, top). After ionizing radiation, the ability of p53 to upregulate p21^{CIP1/WAF1} mRNA was impaired in cells in which APC5 and APC7 gene expression was reduced through RNAi (Fig. 2c, top). The reduction in p21^{CIP1/WAF1} mRNA after APC5 and APC7 knockdown was reflected in the amount of p21^{CIP1/WAF1} protein (Fig. 2c, bottom). Crucially, however, the amount of p53 protein, before and after ionizing radiation, was not dependent on APC5 and/or APC7 gene expression (Fig. 2c, bottom), suggesting that APC5 and APC7 affect p53-regulated transcription of p21^{CIP1/WAF1} directly. Furthermore, knockdown of either APC5 or APC7 had only a modest effect on HSF cell-cycle status, although HSFs did take longer to progress through mitosis when protein expression was reduced by RNAi (Supplementary Fig. S3).

¹Cancer Research UK Institute for Cancer Studies, The Medical School, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. ²Verna & Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas 77030, USA. ³Marie Curie Research Institute, The Chart, Oxted, Surrey RH8 0TL, UK.

⁴Department of Genetics, Center for Genetics and Genomics, and Howard Hughes Medical Institute, Harvard University Medical School, Boston, Massachusetts 02115, USA.

We further investigated the role of APC/C subunits in transcriptional control by chromatin immunoprecipitation (ChIP). After HSFs were treated with ionizing radiation, the amount of p53 recruited to the p21^{CIP1/WAF1} promoter increased (Fig. 2d). CBP/p300 and the APC/C subunits APC3, APC5 and APC7 (Fig. 2d) were also associated with p21^{CIP1/WAF1} promoter elements; their recruitment to this promoter, however, was not considerably altered after ionizing radiation (Fig. 2d, e). Intriguingly, knockdown of APC5, and particularly APC7, markedly reduced the amount of acetylated histone H4 associated with the p21^{CIP1/WAF1} promoter both before and after ionizing radiation (Fig. 2e, top); CBP/p300 association with the p21^{CIP1/WAF1} promoter was, however, unaffected by APC5 and/or APC7 gene knockdown (Fig. 2e, bottom).

We next investigated whether the APC/C affected autoacetylation of lysine 1499 (K1499; ref. 23), and hence the intrinsic acetyltransferase activity, of p300. *In vitro* assays using purified p300 and APC/C (Supplementary Fig. S4a, b) showed that the APC/C enhanced p300 acetyltransferase activity in an ubiquitination-independent manner (Fig. 2f). Consistent with a requirement for APC5 and APC7 in p300 activation *in vivo*, ablation of APC5 and/or APC7 gene expression by RNAi in A549 cells reduced K1499 acetylation in p300 markedly (Fig. 2g), despite an enhancement in p300 protein (Fig. 2g). Amounts of CBP and p300 protein were similarly increased after short interfering (siRNA)-mediated ablation of APC5, APC7 or Cdh1 gene expression in HSFs (Supplementary Fig. S4c, d), suggesting that CBP and p300 might also be substrates for Cdh1-APC/C complexes. It is possible that, akin to transcription factors such as Myc²⁴, CBP/p300

coactivator function might ultimately be controlled by ubiquitin-mediated proteolysis.

Because CBP and p300 function as universal coactivators, we investigated whether the APC/C could similarly regulate CBP/p300-dependent activation of the E2F-1-DP-1 complex²¹. Notably, APC5 and APC7 enhanced E2F-1-DP-1-dependent transactivation in a CBP/p300-dependent manner (Supplementary Fig. S5a). Moreover, ChIP analyses showed that the APC/C subunits APC5, APC7 and APC8, in addition to E2F-1 and CBP/p300, were associated with a region of the *cdc6* promoter encompassing two E2F-1-binding sites (Supplementary Fig. S5b, top). Similar to the p21^{CIP1/WAF1} promoter, the acetylation status of histone H4 associated with the *cdc6* promoter was reduced appreciably *in vivo* (Supplementary Fig. S5b, middle) after APC5 and APC7 knockdown by RNAi (Supplementary Fig. S5c). In agreement with this, overexpression of APC5 and APC7 enhanced the CBP-dependent acetylation of promoter-bound histone H4 (Supplementary Fig. S5b, bottom). ChIP analyses after RNAi also showed that recruitment of APC8, and hence the APC/C holoenzyme, to the *cdc6* promoter was dependent on APC5 and APC7 (Supplementary Fig. S5d, e) and that recruitment of these APC/C components to the *cdc6* promoter was largely dependent on CBP/p300 expression (Supplementary Fig. S5f, g). Furthermore, deconvolution microscopy indicated that both APC5 and APC7 were colocalized with CBP in interphase cells (Supplementary Fig. S6) and that APC5 also colocalized with acetylated-histone H3, and hence active transcription sites, *in vivo* (Supplementary Fig. S6). Together, these

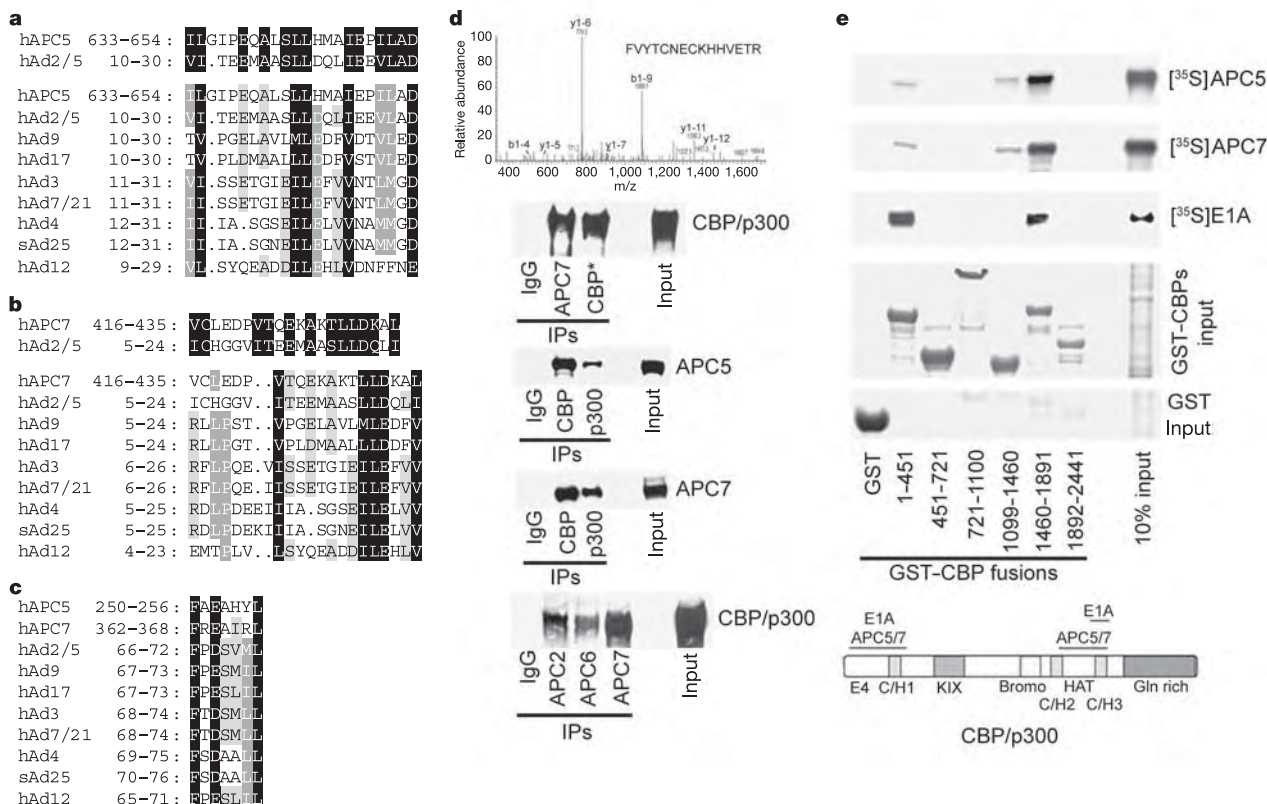


Figure 1 | Binding of APC5, APC7 and E1A to CBP/p300 through conserved domains. **a–c.** Comparative sequence alignment of human APC5 and APC7 with adenovirus E1A. Alignments were done with CLUSTAL W and imported into GeneDoc for shading of conserved residues. **d.** APC5 and APC7 bind CBP/p300 *in vivo* in A549 cells. The upper panel shows mass spectrometric determination of p300 (1,666–1,680) and CBP (1,703–1,717) peptides identified in anti-APC5 immunoprecipitates. The lower four blots show, from top to bottom, coimmunoprecipitation of CBP/p300 with APC7,

APC5 and APC7 with CBP and p300, and CBP/p300 with APC2 and APC6. CBP* immunoprecipitates both CBP and p300. **e.** Top, GST–CBP binding to [³⁵S]- α -methionine-labelled wild-type APC5, APC7 and E1A. Bottom, binding sites for APC5, APC7 and E1A on CBP/p300. E4, ubiquitin E4 ligase activity; C/H1–C/H3, cysteine/histidine-rich regions; KIX, CREB-binding domain; Bromo, bromodomain; HAT, histone-directed acetyltransferase activity; Gln rich, glutamine-rich region.

data support the view that the APC/C functions *in vivo* to regulate transcription.

Targeting of CBP/p300 by E1A is a key event in the transformation process; E1A mutants unable to bind CBP/p300 are defective in transformation^{25–27}. We therefore examined the requirement for APC5 and APC7 in this process. Notably, co-transfection of hooded Lister rat embryo fibroblasts (HLREFs) with either APC5 or APC7 expression plasmids reduced the ability of E1A and *ras* to induce transformation (Fig. 3). Similarly, APC5 suppressed E1A/E1B-mediated transformation, suggesting that it was targeting E1A, and not Ras (Fig. 3a). Consistent with the notion that APC5 and APC7 suppress E1A-mediated transformation by targeting CBP/p300, APC5 and APC7 mutants unable to bind these proteins (Supplementary Figs S2 and S7a) were severely restricted in their ability to suppress E1A-mediated transformation (Fig. 3). GST pull-downs assays confirmed that both APC5 and APC7 competed with E1A for binding to CBP (Supplementary Fig. S7b, c).

These data also suggested that E1A-mediated transformation might be dependent on the functional inactivation of APC/C–CBP/p300 complexes. If so, then a transformation-defective E1A

mutant that is unable to bind CBP/p300, such as RG2 E1A, should regain transforming potential in the absence of APC5 and/or APC7 gene function. Indeed, the ability of RG2 E1A, but not wild-type E1A, to mediate transformation was considerably enhanced in HLREFs in which APC5, APC7, or APC5 and APC7 levels were reduced through RNAi (Supplementary Fig. S8). The different efficacies of RG2 E1A-mediated transformation in these circumstances related directly to the efficiency of gene knockdown (Supplementary Fig. S8). Together, these data indicates that E1A may target APC/C function directly during the transformation process by binding to CBP/p300.

Given that the APC/C regulates cell-cycle progression through targeted ubiquitin-mediated proteolysis, we investigated whether CBP and p300 regulated APC/C function. Like APC3, APC5 and APC7 immunocomplexes, both CBP and p300 immunocomplexes supported the polyubiquitination of cyclin B1 (Fig. 4a), indicating that these proteins were associated with active APC/C ubiquitin ligase complexes. Indeed, both CBP and p300 were also associated *in vivo* with the APC/C activator proteins Cdc20 and Cdh1 (Fig. 4b). To investigate further the roles of CBP and p300 in cell-cycle progression, we ablated CBP and p300 expression by RNAi

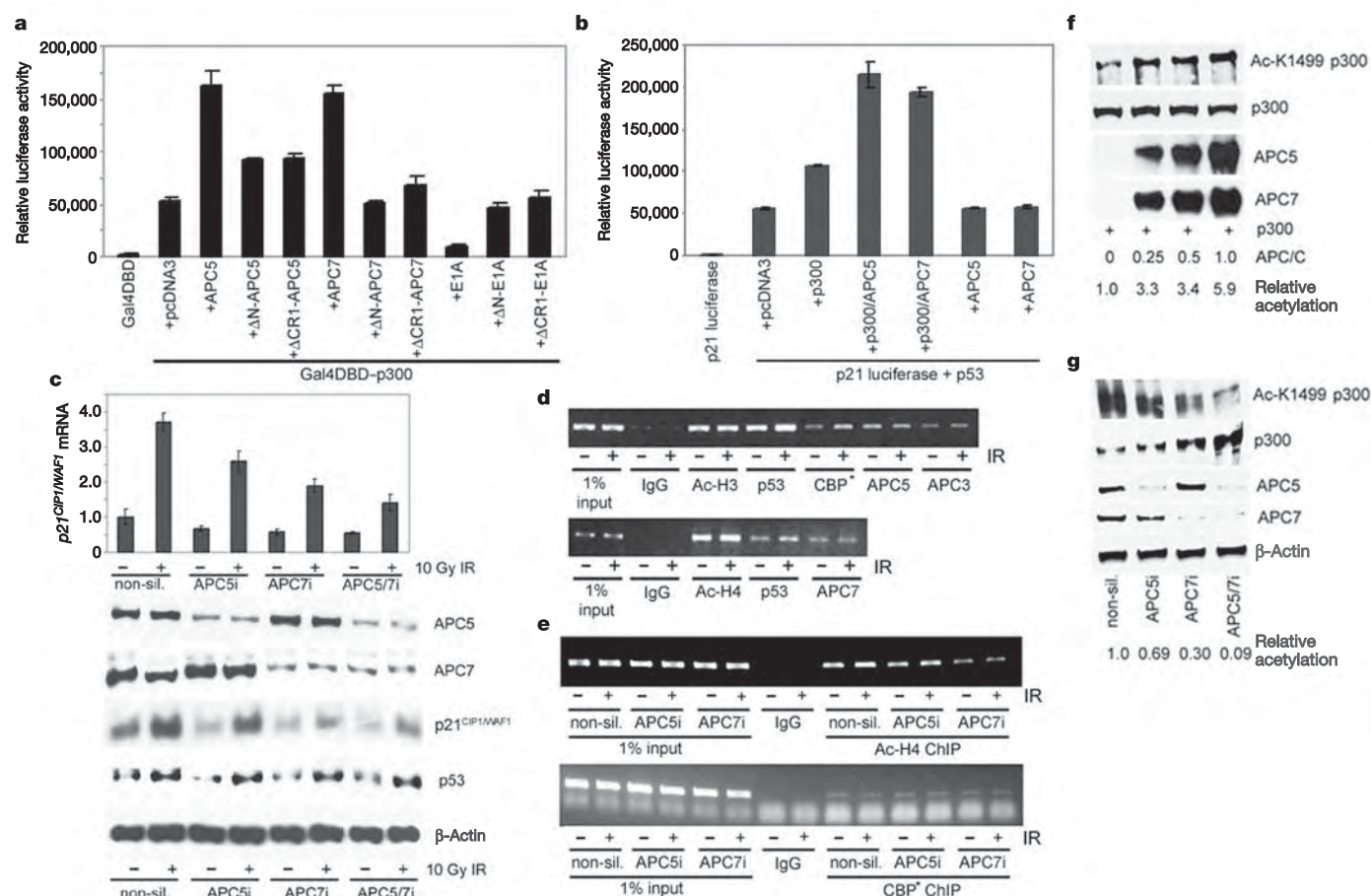


Figure 2 | Role of the APC/C in CBP/p300-regulated transcription.

a, b, APC5 and APC7 potentiate CBP/p300-dependent transcription. HCT116 (**a**) and H1299 (**b**) cells were transfected with the appropriate plasmids, and luciferase and β-galactosidase activities were measured 48 h after transfection. DBD, DNA-binding domain. Data are the mean ± s.d. of 3–6 independent experiments in **a** and two independent experiments in **b**. **c–e**, *In vivo* requirement for APC5 and APC7 in transcription. In **c** and **e**, HSFs were treated with non-silencing RNA (AATTCTCCGACGTGTCACG TTT), or RNA complementary to APC5 (APC5i; AACCTCCGTGTCCAAGA TGT TTT) and/or APC7 (APC7i; AACAGGCACAGATGTTGGATCTT) and 72-h later were then treated with 10 Gy of ionizing radiation (IR). **c**, Four hours after ionizing radiation, RNA and protein were extracted. Top panel

shows p21^{CIP1/WAF1} mRNA (data are the mean ± s.d. of two independent experiments). Lower panels show APC5, APC7, p21^{CIP1/WAF1}, p53 and β-actin proteins. **d, e**, HSFs were subjected to ionizing radiation; 4-h later, protein–DNA complexes were immunoprecipitated with the appropriate antibodies and p21 promoter elements were detected by PCR. CBP* immunoprecipitates both CBP and p300. **f**, *In vitro* acetyltransferase assays, using purified His₆-p300 (200 ng) and different relative amounts of immunopurified APC/C (0, no APC/C; 0.25, 50 ng of APC/C; 0.5, 100 ng of APC/C; 1.0, 200 ng of APC/C) in the presence of 10 μM acetyl-coenzyme A. **g**, Acetylation status of p300 after APC5 and/or APC7 knockdown in A549 cells. p300 immunoprecipitates were subjected to western blot analysis for the detection of acetylated and non-acetylated p300.

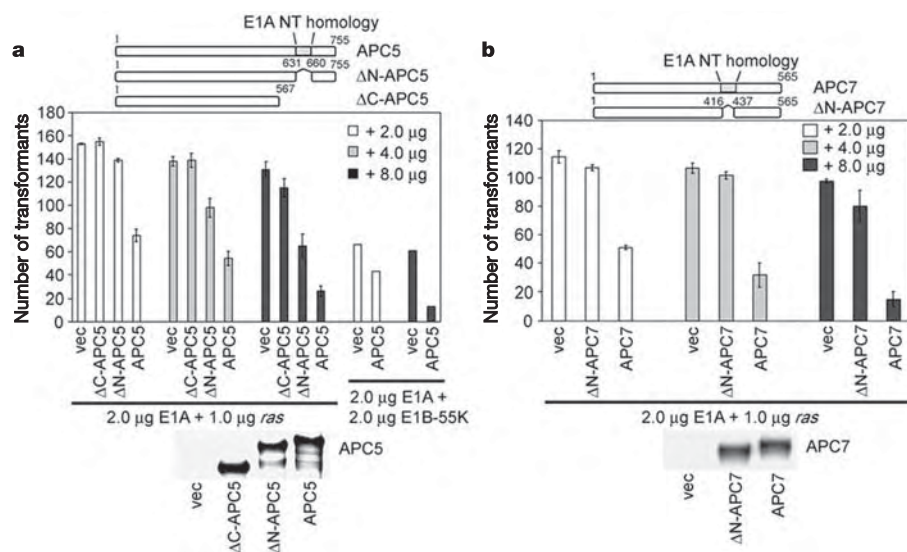


Figure 3 | Suppression of E1A-mediated transformation by APC5 and APC7. HLREF cells were transfected with E1A and activated N-ras, or E1A and E1B-55K, plus the indicated control (vec), APC5 (**a**) and APC7 (**b**) expression plasmids. Data are the mean $\pm$ s.d. of three independent experiments. Western blot analysis shows the expression of exogenously expressed APC5 and APC7 in HLREF cells. The schematic illustrations show the protein products produced by the APC5 (**a**, top) and APC7 (**b**, top) expression plasmids.

(Fig. 4c). Notably, we observed increases in the protein quantities of the known APC/C substrates cyclin B1 and Plk-1 (Fig. 4c), and in the amount of histone H3 phosphorylated on Ser 10 (Fig. 4c) after CBP knockdown. Concordant with this, CBP knockdown led to an

accumulation of cells in mitosis (Fig. 4d and Supplementary Table S1). p300 knockdown, however, did not affect either cyclin B1 or Plk-1 protein (Fig. 4c) or cell-cycle status (Fig. 4d). Furthermore, deconvolution microscopy showed that CBP colocalized with

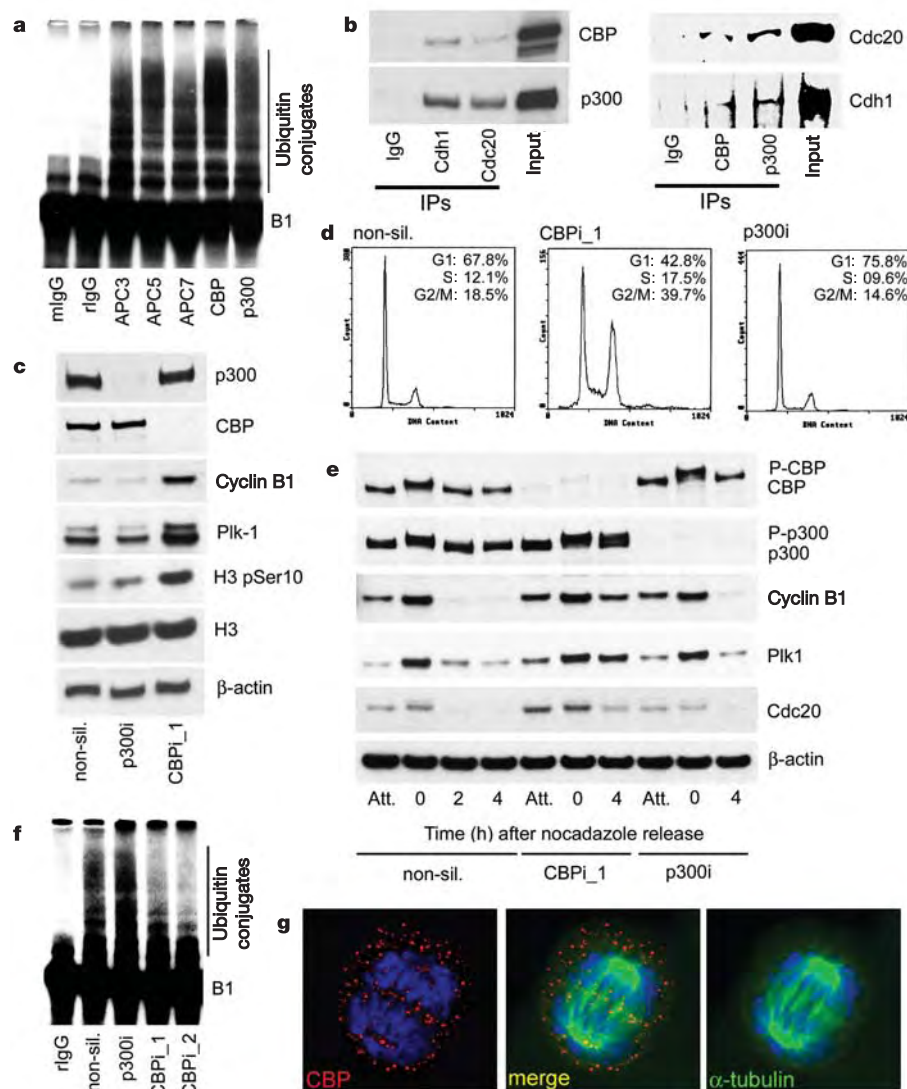


Figure 4 | Role of CBP in mitosis. **a**, Association of CBP and p300 with active APC/C ubiquitin ligase complexes. **b**, *In vivo* association of CBP and p300 with Cdh1 and Cdc20 in U2OS cells. **c**, **d**, Effect of CBP and p300 gene ablation on the quantities of APC/C substrate proteins and on cell-cycle status in U2OS cells. Cells were transfected with either non-silencing RNA or an siRNA complementary to CBP (CBPi_1; AATCAACTCCTGTGTCGTCTTTT) or p300 (p300i; AAGTTCAAACGCCGAGTCTTCTT) and 72-h later were then subjected to either SDS-PAGE and western blot analysis or flow cytometry after staining with propidium iodide. **e**, Mitotic U2OS cells from CBP knockdown cells show a mitotic exit defect. 'Att.' indicates cells remaining attached to the dish after 'mitotic shake-off' with nocadazole ($10 \mu\text{g ml}^{-1}$). **f**, Effect of CBP and p300 gene knockdown on APC/C ubiquitin ligase activity in A549 cells (CBPi_2; AATCCACAGTACCGAGAAATGTT). **g**, Association of CBP with spindle apparatus during anaphase in A549 cells. CBP, red; α -tubulin, green; DAPI, blue.

the APC/C during mitosis (Supplementary Fig. S9). APC5 and CBP colocalized at the spindle poles during metaphase and at the spindle poles and the mid-body during anaphase (Supplementary Fig. S9).

To establish that CBP is required for mitotic progression, we initially treated cells with either non-silencing CBP or p300 siRNA (Fig. 4e). We then arrested these cells in mitosis, using the spindle assembly inhibitor nocadazole, and subsequently released the mitotic 'shake-off' cells back into cycle by withdrawing nocadazole treatment. Mitotic cells collected after CBP knockdown took considerably longer to exit mitosis after release from the nocadazole block, as determined by the amounts of cyclin B1 and Plk-1 (Fig. 4e), than did mitotic cells treated with non-silencing or p300 siRNA. These data suggested that APC/C function was impaired after CBP knockdown. To determine whether ablation of CBP gene expression specifically affected APC/C function, we assayed APC/C ligase activity from cells after RNAi directed against CBP and p300. Notably, APC/C ubiquitin ligase activity was markedly reduced after ablation of CBP, but not p300, gene expression (Fig. 4f). CBP thus enhances APC/C ligase activity, presumably through its inherent E4 ligase capacity²⁸. The idea that CBP is required for mitosis was strengthened by the observation that it colocalized with both mitotic spindles and the spindle poles in anaphase (Fig. 4g).

Taken together, our results suggest that the functional synergy shown by APC/C–CBP/p300 complexes is mediated through the control of acetylation and ubiquitination (Supplementary Fig. S10). The ability of CBP and p300 to regulate diverse cellular functions through interaction with numerous transcription factors and tumour suppressor gene products²⁹ suggests that the APC/C may also be involved in these processes. It is also possible that the APC/C might be acetylated by CBP/p300 and this might have a positive role in APC/C function. The observation that both APC5 and APC7 interact independently with CBP and p300 indicates a level of complexity that may define discrete biological properties for different complexes containing APC/C–CBP and APC/C–p300.

METHODS

Immunoprecipitation. For immunoprecipitation, cells were lysed in buffer containing 50 mM Tris-HCl (pH 7.4), 0.825 M NaCl and 1% Nonidet P-40, sonicated and cleared by centrifugation. Immunoprecipitation, with the appropriate antibodies, was done as described¹⁰.

GST pull-downs assays. Typically, 10 µg of the appropriate GST fusion protein was mixed with *in vitro* translated protein (Promega) labelled with [³⁵S]L-α-methionine (Amersham). Pull-downs assays were done as described¹⁰. [³⁵S]proteins were visualized by fluorography with Amplify (Amersham) and autoradiography of dried polyacrylamide gels.

Transfection and reporter assays. Complexes of DNA and LipofectAmine (Invitrogen) were incubated with cells for 6 h according to the manufacturer's instructions. At the appropriate times after transfection, the cells were lysed, and luciferase and β-galactosidase activities were determined using reagents and protocols supplied by the manufacturer (Promega).

Mass spectrometry. Immunoprecipitated proteins were reduced, carboxy-methylated and digested with trypsin (Promega). Tryptic peptides were separated on a C₁₈-reverse phase PepMap HPLC column and the eluate was sprayed into a ThermoFinnigan LCQDecaXP IonTrap mass spectrometer by a 20-µm Picotip emitter, using 2.5 kV at 155 °C. Intense peaks were analysed by tandem mass spectrometry using 35% collision energy. The collision induced dissociation spectra were analysed against the human non-redundant protein database using Sequest (Flicka).

Real-time PCR. Total RNA was extracted with Trizol reagent (Invitrogen) and reverse-transcribed with avian myeloblastosis virus reverse transcriptase (Promega). Expression of specific mRNAs was determined with the ABI PRISM 7700 sequence detection system. We used the following p21^{CIP1/WAF1} primers: forward, 5'-gcagaccagcatgacagatttc-3'; reverse, 5'-ggattaggctctcttggga-3'; probe (FAM), 5'-ccactccaaacgcccgtgatctt-3'. Reactions were multiplexed with a control probe (VIC) for 18S ribosomal RNA (PE Biosystems).

Transformation. Primary HLREFs were prepared from 18-day-old embryos and used at passage two. For each experimental condition, plasmid DNA, made up to 15 µg with heat-denatured salmon sperm DNA, was added to 2 × 10⁶ cells in a final volume of 250 µl of medium. Cells were electroporated using a Bio-Rad Gene Pulsar at 960 µF and 220 V in 4-mm cuvettes. After transfection, cultures

were fed every 3 d. G418 selection (200 µg ml⁻¹ final concentration) was initiated 48 h after electroporation and continued until day 14, when transformed foci were counted under low-power microscopy. Where appropriate, siRNAs and vectors expressing short hairpin RNAs were included in the transformation assay.

Immunofluorescence. Cells were fixed in paraformaldehyde (4% (w/v) in PBS) and permeabilized with acetone. After immunostaining, cells were viewed with a Nikon Eclipse E600 microscope. z-Layer images (0.2-µm horizontal sections) were recorded with an Openlab v3.0.9. Biovision software package (Improvision) and selected z-layers deconvolved. Single, equivalent z-layers were merged to show colocalization.

APC/C ligase assays. Active APC/C complexes were immunoprecipitated from cells as described³⁰ with Sepharose conjugated to Protein G. We assayed 4 µl of packed Protein-G-Sepharose immunocomplexes for APC/C ubiquitin ligase activity in a final volume of 10 µl containing 50 mM Tris (pH 7.5), 5 mM MgCl₂, 5 mM ATP, 10 mM creatine phosphate, 350 U ml⁻¹ of creatine phosphokinase, 1 mg ml⁻¹ of ubiquitin, 14.4 µg ml⁻¹ of ubiquitin aldehyde, 35 µg ml⁻¹ of rabbit E1, 40 µg ml⁻¹ of UbcH10, 40 µg ml⁻¹ of UbcH4, 40 µg ml⁻¹ of UbcH5 and 1 µl of [³⁵S]cyclin B1. Samples were incubated for 30 min at 37 °C and resolved by SDS-PAGE, and the dried gels were visualized by fluorography.

Chromatin immunoprecipitation. After the appropriate treatment cells were crosslinked with 1% (v/v) formaldehyde. Chromatin immunoprecipitation was done using lysis buffers and protocols supplied by the manufacturer (Upstate Biotechnology).

Flow cytometry. Cells were fixed at -20 °C in 70% (v/v) ethanol for 1 h, before treatment with RNase A (10 µg ml⁻¹) and pepsin (2 mg ml⁻¹) at 37 °C for an additional 1 h. Cells were resuspended in PBS containing propidium iodide (5 µg ml⁻¹) and subjected to flow cytometry using a Coulter Epics XL-MCL instrument.

Received 5 July; accepted 17 August 2005.

- Harper, J. W., Burton, J. L. & Solomon, M. J. The anaphase-promoting complex: it's not just for mitosis anymore. *Genes Dev.* **16**, 2179–2206 (2002).
- Peters, J.-M. The anaphase-promoting complex: proteolysis in mitosis and beyond. *Mol. Cell* **9**, 931–943 (2002).
- Bashir, T., Dorrello, N. V., Amador, V., Guardavaccaro, D. & Pagano, M. Control of the SCF^{Skp2-Cks1} ubiquitin ligase by the APC/C^{cdh1} ubiquitin ligase. *Nature* **428**, 190–193 (2004).
- Wei, W. *et al.* Degradation of the SCF component Skp2 in cell-cycle phase G1 by the anaphase-promoting complex. *Nature* **428**, 194–198 (2004).
- Eckner, R. *et al.* Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. *Genes Dev.* **8**, 869–884 (1994).
- Lundblad, J. R., Kwok, R. P. S., Lurance, M. E., Harter, M. L. & Goodman, R. H. Adenoviral E1A-associated protein p300 as a functional homologue of the transcriptional co-activator CBP. *Nature* **374**, 85–88 (1995).
- O'Connor, M. J., Zimmermann, H., Nielsen, S., Bernard, H.-U. & Kouzarides, T. Characterization of an E1A–CBP interaction defines a novel transcriptional adapter motif (TRAM) in CBP/p300. *J. Virol.* **73**, 3574–3581 (1999).
- Gallimore, P. H. & Turnell, A. S. Adenovirus E1A: remodelling the host cell, a life or death experience. *Oncogene* **20**, 7824–7835 (2001).
- Vodermaier, H. C., Gieffers, C., Maurer-Stroh, S., Eisenhaber, F. & Peters, J.-M. TPR subunits of the anaphase-promoting complex mediate binding to the activator protein CDH1. *Curr. Biol.* **13**, 1459–1468 (2003).
- Turnell, A. S. *et al.* Regulation of the 26S proteasome by adenovirus E1A. *EMBO J.* **19**, 4759–4773 (2000).
- Zhang, X. *et al.* The targeting of the proteasomal regulatory subunit S2 by adenovirus E1A causes inhibition of proteasomal activity and increased p53 expression. *J. Biol. Chem.* **279**, 25122–25133 (2004).
- Kwok, R. P. S. *et al.* Nuclear protein CBP is a coactivator for the transcription factor CREB. *Nature* **370**, 223–226 (1994).
- Bannister, A. J. & Kouzarides, T. The CBP coactivator is a histone acetyltransferase. *Nature* **384**, 641–643 (1996).
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
- Yuan, W., Condorelli, G., Caruso, M., Felasani, A. & Giordano, A. Human p300 protein is a coactivator for the transcription factor MyoD. *J. Biol. Chem.* **271**, 9009–9013 (1996).
- Chakravarti, D. *et al.* Role of CBP/p300 in nuclear receptor signalling. *Nature* **383**, 99–103 (1996).
- Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 403–414 (1996).
- Gu, W., Shi, X. L. & Roeder, R. G. Synergistic activation of transcription by CBP and p53. *Nature* **387**, 819–823 (1997).
- Lill, N. L., Grossman, S. R., Ginsberg, D., DeCaprio, J. & Livingston, D. M. Binding and modulation of p53 by p300/CBP coactivators. *Nature* **387**, 823–827 (1997).

20. Avantiaggiati, M. L. *et al.* Recruitment of p300/CBP in p53-dependent signal pathways. *Cell* **89**, 1175–1184 (1997).
21. Trouche, D., Cook, A. & Kouzarides, T. The CBP co-activator stimulates E2F1/DP1 activity. *Nucleic Acids Res.* **24**, 4139–4145 (1996).
22. El-Deiry, W. S. *et al.* WAF1, a potential mediator of p53 tumour suppression. *Cell* **75**, 817–825 (1993).
23. Thompson, P. R. *et al.* Regulation of the p300 HAT domain via a novel activation loop. *Nature Struct. Mol. Biol.* **11**, 308–315 (2004).
24. Salghetti, S. E., Kim, S. Y. & Tansey, W. P. Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc. *EMBO J.* **18**, 717–726 (1999).
25. Lillie, J. W., Green, M. & Green, M. R. An adenovirus E1a protein region required for transformation and transcriptional repression. *Cell* **46**, 1043–1045 (1986).
26. Jelsma, T. N. *et al.* Sequences in E1A proteins of human adenovirus 5 required for cell transformation, repression of a transcriptional enhancer, and induction of proliferating cell nuclear antigen. *Virology* **170**, 120–130 (1989).
27. Wang, H. G. H. *et al.* Identification of specific adenovirus-E1A N-terminal residues critical to the binding of cellular proteins and to the control of cell growth. *J. Virol.* **67**, 476–488 (1993).
28. Grossmann, S. R. *et al.* Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science* **300**, 342–344 (2003).
29. Goodman, R. H. & Smolik, S. CBP/p300 in cell growth, transformation and development. *Genes Dev.* **14**, 1553–1577 (2000).
30. Kramer, E. R., Gieffers, C., Holz, G., Hengstschlager, M. & Peters, J. M. Activation of the human anaphase-promoting complex by proteins of the CDC20/Fizzy family. *Curr. Biol.* **8**, 1207–1210 (1998).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Brockmann, M. J. Campbell, R. T. Hay, D. P. Lane, M. W. Kirschner, T. Kouzarides, W. L. Kraus, E. Moran, J. S. Mymryk and J. Pines for reagents; C. J. McCabe for help with PCR analysis; and J.S.M. for help with alignments and for critically reading the manuscript. This work was supported by a Cancer Research UK programme grant. P.H.G. is a Cancer Research UK Gibb Fellow. G.S.S. was supported by a Cancer Research UK career development fellowship and an EMBO long-term fellowship. H.Y. was supported by Marie Curie Cancer Care and the Association for International Cancer Research. S.J.E. is an investigator with the Howard Hughes Medical Institute and is supported by a grant from the NIH.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to A.S.T. (A.S.Turnell@bham.ac.uk).

CORRIGENDUM

doi:10.1038/nature04362

A network-based analysis of systemic inflammation in humans

Steve E. Calvano, Wenzhong Xiao, Daniel R. Richards, Ramon M. Felciano, Henry V. Baker, Raymond J. Cho, Richard O. Chen, Bernard H. Brownstein, J. Perren Cobb, S. Kevin Tschoeke, Carol Miller-Graziano, Lyle L. Moldawer, Michael N. Mindrinos, Ronald W. Davis, Ronald G. Tompkins, Stephen F. Lowry & the Inflammation and Host Response to Injury Large Scale Collaborative Research Program†

Nature 437, 1032–1037 (2005) doi:10.1038/nature03985

In this Letter, the affiliations of authors participating in the Inflammation and Host Response to Injury Large Scale Collaborative Research Program are incorrectly listed. The renumbered and amended footnote listing is given here.

†**Inflammation and Host Response to Injury Large Scale Collaborative Research Program** Paul E. Bankey¹, Timothy R. Billiar², David G. Camp³, George Casella⁴, Irshad H. Chaudry⁵, Mashkoor A. Choudhry⁵, Charles Cooper⁶, Asit De¹, Constance Elson⁷, Bradley Freeman⁸, Richard L. Gamelli⁹, Celeste Campbell-Finnerty¹⁰, Nicole S. Gibran¹¹, Douglas L. Hayden⁷, Brian G. Harbrecht², David N. Herndon¹⁰, Jureta W. Horton¹², William J. Hubbard⁵, John L. Hunt¹³, Jeffrey Johnson¹⁴, Matthew B. Klein¹⁵, James A. Lederer¹⁶, Tanya Logvinenko⁷, Ronald V. Maier¹¹, John A. Mannick¹⁶, Philip H. Mason⁶, Bruce A. McKinley¹⁷, Joseph P. Minei¹², Ernest E. Moore¹⁴, Frederick A. Moore¹⁷, Avery B. Nathens¹¹, Grant E. O'Keefe¹¹, Laurence G. Rahme¹⁸, Daniel G. Remick¹⁹, David A. Schoenfeld⁷, Martin G. Schwacha⁵, Michael B. Shapiro²⁰, Geoffrey M. Silver⁹, Richard D. Smith³, John D. Storey²¹, Mehmet Toner²², H. Shaw Warren²³ & Michael A. West²⁰

Affiliations for participants: ¹Department of Surgery, University of Rochester School of Medicine, Rochester, New York 14642, USA. ²Department of Surgery, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15213, USA. ³Pacific Northwest National Laboratory, Richland, Washington 99352, USA. ⁴Department of Statistics, University of Florida, Gainesville, Florida 32611, USA. ⁵Department of Surgery, University of Alabama School of Medicine, Birmingham, Alabama 35294, USA. ⁶Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School, Cambridge, Massachusetts 02139, USA. ⁷Department of Biostatistics, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. ⁸Department of Surgery, Washington University School of Medicine, St. Louis, Missouri 63110, USA. ⁹Department of Surgery, Loyola University Stritch School of Medicine, Maywood, Illinois 60153, USA. ¹⁰Department of Surgery, University of Texas Medical Branch, Shriners Burns Hospital, Galveston, Texas 77550, USA. ¹¹Department of Surgery, University of Washington Harborview Medical Center, Seattle, Washington 98104, USA. ¹²Department of Surgery, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ¹³Division of Trauma, Burns, and Critical Care, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ¹⁴Department of Surgery, University of Colorado Denver Health Medical Center, Denver, Colorado 80204, USA. ¹⁵Burn Center and Division of Plastic Surgery, University of Washington Harborview Medical Center, Seattle, Washington 98104, USA. ¹⁶Department of Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. ¹⁷Department of Surgery, University of Texas Houston Health Science Center, Houston Medical School, Houston, Texas 77030, USA. ¹⁸Department of Molecular Biology, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. ¹⁹Department of Medical Science, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. ²⁰Department of Surgery, Northwestern University Medical School, Chicago, Illinois 60611, USA. ²¹Department of Biostatistics, University of Washington, Seattle, Washington 98195, USA. ²²Center for Engineering in Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. ²³Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02129, USA.

CORRIGENDUM

doi:10.1038/nature04363

DNA sequence and analysis of human chromosome 18

Chad Nusbaum, Michael C. Zody, Mark L. Borowsky, Michael Kamal, Chinnappa D. Kodira, Todd D. Taylor, Charles A. Whittaker, Jean L. Chang, Christina A. Cuomo, Ken Dewar, Michael G. Fitzgerald, Xiaoping Yang, Amr Abouelleil, Nicole R. Allen, Scott Anderson, Toby Bloom, Boris Bugalter, Jonathan Butler, April Cook, David DeCaprio, Reinhard Engels, Manuel Garber, Andreas Gnirke, Nabil Hafez, Jennifer L. Hall, Catherine Hosage Norman, Takehiko Itoh, David B. Jaffe, Yoko Kuroki, Jessica Lehoczy, Annie Lui, Pendexter Macdonald, Evan Mauceli, Tarjei S. Mikkelsen, Jerome W. Naylor, Robert Nicol, Cindy Nguyen, Hideki Noguchi, Sinéad B. O'Leary, Keith O'Neill, Bruno Piqani, Cherylyn L. Smith, Jessica A. Talamas, Kerri Topham, Yasushi Totoki, Atsushi Toyoda, Hester M. Wain, Sarah K. Young, Qiangdong Zeng, Andrew R. Zimmer, Asao Fujiyama, Masahira Hattori, Bruce W. Birren, Yoshiyuki Sakaki & Eric S. Lander

Nature 437, 551–555 (2005) doi:10.1038/nature03983

The name of Keith O'Neill was accidentally omitted from the published author list. He is at the first affiliation in the address list.

ERRATUM

doi:10.1038/nature04361

Astronomical pacing of methane release in the Early Jurassic period

David B. Kemp, Angela L. Coe, Anthony S. Cohen & Lorenz Schwark

Nature 437, 396–399 (2005)

In the labelling of Fig. 1 of this Letter, the spelling of '*D. semicelatum*' was accidentally reversed to read '*D. mutalecimes*'. It appears correctly in the text.

Ion channels and stem cells

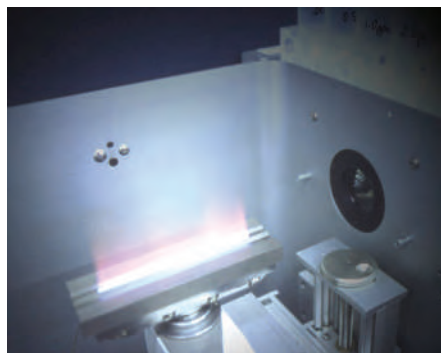
Ion channels, stem cells and cell signalling are the focus of intense interest in both cell biology and drug discovery. Pete Moore takes a look at what's on offer for the researcher.

AURORA BIOMED

Ion channels act as electrical gatekeepers in cell membranes, and are responsible for the generation and propagation of nerve impulses, muscle contraction, and many other biological processes. With more than 400 ion-channel genes identified in the human genome, interest in detecting and measuring their activity is burgeoning.

A high-throughput method of assessing the function of outward-rectifying potassium channels is to monitor the flow of tracer ions through them. In the case of potassium channels, rubidium ions (Rb^+) are used as a tracer because Rb^+ has similar characteristics to K^+ but is not present in biological systems and so there is no background noise. Trace amounts of Rb^+ (as low as 0.05 mg l^{-1}) can be detected using flame atomic absorption spectroscopy with the Ion Channel Reader (ICR) from Aurora Biomed of Vancouver, British Columbia. The ICR can be used to study voltage- and ligand-gated potassium channels as well as sodium channels and chloride channels.

Another way of studying ion-channel activity is to monitor changes in membrane potential. Invitrogen of Carlsbad, California, and PerkinElmer of Boston, Massachusetts, have



Aurora's Ion Channel Reader measures Rb^+ flow.

recently joined forces to offer a combination of Invitrogen's Voltage Sensor Probes ion-channel reagents and PerkinElmer's CellLux Fluorescence Cellular Screening Platform. This assay is based on fluorescence resonance excitation transfer (FRET); it uses a coumarin-phospholipid FRET donor that binds to the exterior of the cell membrane and a negatively charged FRET acceptor. In resting cells the two probes associate with the membrane exterior, resulting in efficient FRET and a red fluorescence signal. When a cell becomes depolarized

as ions flow through channels, the FRET acceptor rapidly translocates to the other membrane face. Exciting the donor probe now generates a blue fluorescence signal.

Tracking channels

Ion-channel localization can affect cell function dramatically, and ChanTest of Cleveland, Ohio, offers antibody-based tests for detecting intracellular ion-channel trafficking. "In cystic fibrosis, 50% of families have a defect that prevents the CFTR channel protein being transported to the cell surface, and for the hereditary form of the hERG disease, about half of the mutations in the hERG channel protein affect trafficking," says ChanTest's chief executive officer Arthur 'Buzz' Brown. Blocking the function of the hERG potassium ion channel in cardiac muscle may be a major adverse drug effect as it can cause arrhythmia and sudden cardiac death, and all new drugs must be tested for whether they block this channel. In ChanTest's HERG-Lite assay, human embryonic kidney (HEK) cells express a version of the hERG channel carrying a hemagglutinin epitope. Protein turnover replenishes hERG channels about every 12 hours, so the cells are incubated overnight with

MAXIMIZING RETURN

CELLECTRICON

Although most patch-clamp technologies seek to maximize the number of cells rushed through the system, Owe Orwar and his colleagues at Cellectricon, a start-up company based in Gothenburg, Sweden, have developed a platform that maximizes the information gained from each cell. The result is a powerful tool for

secondary screening in drug discovery. Their Dynaflow technology uses conventional glass pipette patch clamping, in combination with a novel microfluidic device for controlled delivery of drug solutions.

Solutions of drugs or drug combinations are placed in up

to 48 wells, each of which is connected to a measurement chamber by micrometre-diameter

channels. These solutions can be directed through the chamber with high precision. "Dynaflow uses the unique properties of fluids when they are running at

very low Reynolds numbers. When the fluids come out from a tiny channel in the open volume they behave as if they are still in channels — they do not mix," says Orwar.

With no turbulence, diffusion would be the only chance of mixing

between solution batches, but the timescales used are too short for that to occur. Consequently, Dynaflow can provide step changes in drugs or drug concentrations, with a change every 30 milliseconds if desired. "It is the most precise technology in the world to titrate receptors," claims Orwar. "You can see it as a microfluidic device that generates a barcode of chemicals, and the cell effectively reading the barcode," he adds.

The ability to squeeze so much data out of a single cell enables some users to claim a ten-fold increase in productivity. By using carefully considered combinations of drugs in each well, cells can be taken through physiologically relevant conditions that relate to many different disease states. "In effect, it gives you the option of passing a chemical waveform over the cell while constantly recording from it," says Orwar.

P.M.



Cellectricon's Dynaflow patch-clamp system combines microfluidics and patch-clamping technology.

the test compounds. The next day an antibody for the epitope is added along with a second antibody that produces a chemiluminescent signal. "If you don't permeabilize the membrane you can count the number of channels at the cell surface — it's simple and fast," says Brown. ChanTest's FAST & Lite service runs the antibody-based test alongside automated patch clamping to assess channel function. Assays for other channels are being developed and ChanTest has been awarded a small-business innovation grant from the National Institutes of Health to automate its system.

Patch clamping goes automated

Although indicator-based methods are fast and inexpensive, the gold-standard for assessing ion channels is the Nobel prize-winning technique of patch clamping developed by Erwin Neher and Bert Sakmann in the 1970s. The conventional manual method involves a glass micropipette filled with an ionic solution that electrically connects a silver-silver chloride electrode wire to a small patch of cell membrane. A vital part of the procedure is to get an electrical seal of at least 1 gigaohm between the pipette tip and the membrane; without this seal the tiny currents that pass through the channels in the membrane patch cannot be measured. The drawback is that the technique requires considerable expertise, hours are spent poring over a microscope, and recordings can only be taken from one cell at a time. But over the past few years automation has entered this green-fingered science.

A major player in the automated patch-



IonWorks Quattro from Molecular Devices.

clamp market is Molecular Devices of Sunnyvale, California, which merged last year with imaging specialists Axon Instruments. Molecular Devices has two high-throughput automated patch-clamping systems that can collect between 100 and 2,000 patch-clamping data points a day, depending on configuration.

Both instruments work by sucking cells down against 1–2 μm diameter pores in the base of multi-well plates. The PatchXpress 7000A uses 16-well, glass SealChip plates made by Aviva Biosciences of San Diego, California. The machine places cells in each well and suction holds one cell that falls on the pore in place with sufficient strength to create an electrical seal of 1 gigaohm. The machine uses suction to disrupt the cell membrane to access the interior

of the cell, and currents are measured across the entire cell surface. "You are, in effect, reversing traditional patch clamping by having the ground electrode measuring from the inside of the cell rather than from the outside," says Steve Davenport, vice-president of Europe for Molecular Devices. Each well is controlled and monitored individually and cells can be sealed for 30 minutes or more — during which time test compounds can be added to and flushed from the well. A single run takes around 45 minutes. The PatchXpress platform works well for both voltage-gated and ligand-gated ion channels and yields high-quality data comparable to the conventional manual patch-clamp method.

IonWorks Quattro from the same company uses a 384-well Patch Plate, but wells share electronics. "This makes sense for a screening instrument where you need the highest throughput possible without compromising the pharmacology," explains Davenport. The system uses a new technology developed by Molecular Devices called Population Patch Clamp (PPC). PPC uses 64 holes versus a single hole in each well of the Patch Plate. This enables the signal from up to 64 cells in each well to be averaged. "The advantage of PPC over conventional single-hole planar patch-clamp is the reduction in biological variability and substantial increase in the success rate of obtaining a data point from each measurement," says Davenport. Using IonWorks, scientists can measure up to 2,000 data points per day.

This speed doesn't come cheap. Both

MOLECULAR DEVICES

STEM-CELL OPTIONS

It's easy to focus on the kit and forget the really important part of the system — the cell. Cells of most interest with respect to ion channels include neurons and heart cells, which cannot be grown for long in culture and do not divide.

Many of the cell lines used in ion-channel work are, therefore, stem cells and cell lines engineered to express specific channels. These include human embryonic kidney (HEK293) and Chinese hamster ovary (CHO) lines. bSys of Basel,

Switzerland, offers a wide range of screening techniques, but chief executive officer Daniel Konrad believes that one of the company's chief advantages is their skill in selecting and fine-tuning cells.

"Each clone of cells is subtly different, and only trialling with many different sources can show which expression system is ideal," he says. bSys also works hard to find the right suspension protocol. This can make the difference between cells that generate 200

picoamp currents and those that can generate 500–1,000 picoamps and can be used in robotic screening systems, says Konrad.

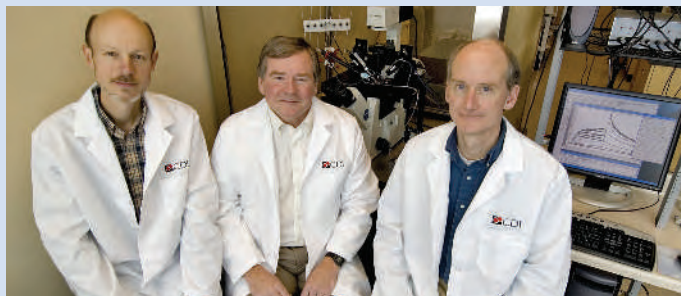
A new company moving into the designer-cell niche is Cellular Dynamics International (CDI) of Madison, Wisconsin, founded by noted human embryonic stem-cell researchers James Thomson, Craig January and Timothy Kamp of the University of Wisconsin. CDI will initially focus on developing HEK cell and cardiomyocyte-based screening services to the pharmaceutical and biotechnology industries, and plans to have a drug-screening service running by the first quarter of 2006.

On the other side of the Atlantic, in Edinburgh, UK, the European arm of Stem Cell Sciences, founded by Peter Mountford in Melbourne, Australia, is developing neural stem (NS) cell lines from the Universities of Edinburgh and Milan. These cells are thought to be

phenotypically similar to the NS cells found *in vivo*. Derived from human and animal embryonic stem (ES) cells and from fetal and adult brain tissue, NS cells have great potential in biomedical research because of their homogeneity, their ability to self-renew indefinitely, and their relative ease of manipulation. Stem Cell Sciences is establishing a service for generating specifically mutated NS cells from engineered ES cells and transgenic animals. NS cells are attractive candidates for *in vitro* drug screening and may also be useful for cellular therapy for conditions such as Parkinson's disease and epilepsy.

R&D Systems of Minneapolis, Minnesota offer ready-to-use primary cortical stem cells derived from rat embryos and the kits to grow them. The cells are validated for differentiation into astrocytes, neurons and oligodendrocytes.

P.M.



Cellular Dynamics International: James Thomson (right), Timothy Kamp (left) and Craig January.

B. FRITZ

NANION

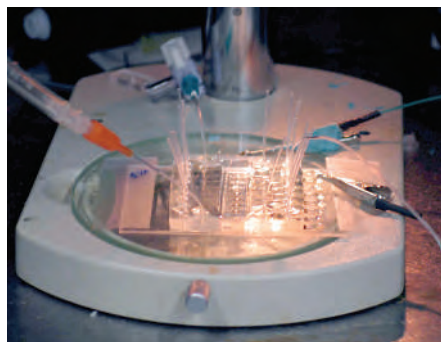
machines cost up to US\$400,000. And although the IonWorks platform works well for voltage-gated channels, where you can adjust the voltage at the same time as recording, it will not work for fast ligand-gated channels, whose currents often last a millisecond or less, as the machine cannot add test compounds and record simultaneously.

Contenders aiming to overcome the ligand-gated channel barrier in automated patch clamping also include Sophion Bioscience of Ballerup, Denmark, which uses a microfluidics approach. Its QPatch 16 operates 16 independent patch-clamp sites, each comprising a flat silicon chip with recording electrodes, a patch-clamp hole, pipetting wells and integrated microfluidic glass flow channels for applying solutions. "QPatch 16 also provides a cell preparation facility in which the cells are suspended in culture medium until right before the experiment. This ensures that cells are kept viable and healthy, and enables unattended operation for at least 4 hours," says Niels Willumsen, a senior executive at Sophion. The integrated microfluidic flow channels of the QPlate allow sequential application of multiple compounds at very low volumes (around 5 μ l) from four to eight pipette tips, and ensure the fast solution exchange (about 50 ms) required to study ligand-gated ion channels. The modular design can be upgraded to a 48-channel system and the machine can give 250–1,200 data points per working day.

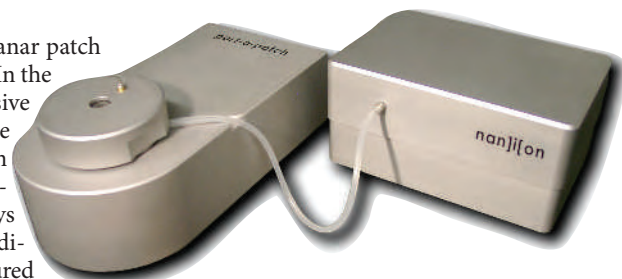
On a smaller scale, Fred Sigworth and Kathryn Klemic at Yale University, New Haven,

Connecticut, have developed a planar patch clamp that can be built in the lab. "In the future, instead of buying an expensive chip, a lab might have a little device that can make an electrode, or an array of little electrodes, by moulding them out of silicon rubber," says Sigworth. A thin layer of polydimethylsiloxane (PDMS) resin is poured on to a plate containing a 2- μ m diameter hole. Before the PDMS cures, air is blown through the hole, creating a 1- μ m hole in the rubber sheet. After peeling the sheet off the plate, exposure of the surface to plasma oxidation creates a 100- μ m thick glassy surface layer of SiO₂. "On the one hand you have a hydrophobic silicone rubber base, then you create this thin layer of glass that the cell rests on — to a cell it looks a lot like a conventional glass electrode," says Klemic.

In expert hands, the best systems for patch



Do-it-yourself: the PDMS microfluidic patch-clamp system in use.



Nanion's Port-a-Patch makes patch clamping easy for the novice.

clamping can currently detect a pulse of about 150 elementary charges: equivalent to a flow of 150 sodium ions. "The grand challenge would be to resolve single elementary charges. Then you could watch a lot of really interesting processes such as the turnover of ions in pumps," says Sigworth. He is unsure whether this single-ion resolution will ever be possible, but thinks that it may be possible to mould the PDMS sufficiently carefully to reduce the capacitance in the system and substantially increase the resolution.

Sigworth is also intrigued by the Port-a-Patch system developed by Nanion Technologies, a spin-off from the Centre for Nanoscience at the University of Munich in Germany. The beauty of Port-a-Patch is its ease of use. "It's basically a bench-top patch clamp. You pipette in the cells, close the lid and make the recording," he says. Nanion claims that this turn-key solution only takes half an hour to set up. "We run one-day training courses, and the system is easily used by people who have no experience in electrophysiology," says Nanion's

F. SIGWORTH & K. KLEMIC

BANKING ON STEM CELLS

Human stem cells are valuable commodities: as well as their medical potential, their pristine naivety makes them attractive as gold-standard cell lines for research. Stem-cell banks, where owners deposit their precious products and would-be investigators apply for loans, are now being developed.

The most advanced is the UK Stem Cell Bank, based at the National Institute for Biological Standards and Controls in Potters Bar, near London. Initiated in September 2002, and funded by the Medical Research Council and the Biotechnology and Biological Sciences Research Council since January 2003, it has the aim of providing a repository for all types of human stem-cell lines.

"As of October 2005, we have 24 stem-cell lines approved for accession into the bank," says director Glyn Stacey, but none is yet ready for sending out to

end-users. That probably won't be until early 2006. "The process is complex. It is not like growing an ordinary cell line where you could create and quality control a bank within a few months of receiving the cells," says Stacey. One time-consuming step is the creation of agreements for depositors and recipients, with each cell type presenting different problems and opportunities. Exploitation will be controlled by the depositor who retains ownership of the cells.

Legal issues aside, stem cells are challenging to grow. The main problem is scaling up to provide hundreds of ampoules of cells at identical passage levels and stages of differentiation. "It could take an entire day for a highly skilled person to dissect and recover cells from just one line," says Stacey. And cultures have to be characterized and checked for contamination before release.

All lines currently in the bank are human embryonic stem (ES) cells. "We have had some contact with people who think they have adult stem cell lines, but they are being careful about characterization," says Stacey.



Glyn Stacey: stem cells are challenging to grow.

A few ampoules of each cell line have been frozen as back-up, whereas the master bank contains 20 or 30 ampoules. The distribution stock may eventually contain around a hundred ampoules of each line. Stacey hopes that early in 2006 the bank's website will start tracking progress of the cell lines that will be available to researchers.

A few other initiatives are taking shape. The US National Stem Cell Bank will be located at the WiCell Research Institute, in Madison, Wisconsin, with a \$16.1 million, four-year National Institutes of Health grant. It will acquire, store, characterize and distribute human embryonic stem-cell lines, but will be limited to those approved for federal funding. After a year of legal wrangling, a stem-cell bank is taking shape at the University of Granada in Spain, and others are being considered in Australia and South Korea.

MRC

P.M.

chief executive officer Niels Fertig. It is the only automated device that addresses the need for low throughput with high accuracy, Fertig claims.

The Port-a-Patch system uses planar borosilicate glass chips (100- μm thick) in which a conical pore of 1- μm diameter is micromachined. The pore has the three-dimensional geometry of an inverted pipette tip and cells are simply positioned via suction. It creates a strong electrical seal with the cell and is ideal for whole-cell patch clamping. Single-channel recordings can be performed in a cell-attached configuration. A software-controlled eight-channel microfluidics add-on can deliver sufficiently rapid changeover of solutions to allow the study of ligand-gated channels. A robotic version of the system that will run 16 patches at a time is in prospect.

The Flyscreen, an alternative approach for moderate-throughput applications from flyion of Tübingen, Germany, can perform 100 to 500 independent whole-cell screens per day. The instrument uses glass micropipettes, into which cells are loaded. As the cells settle, a single one falls towards the tip and wedges near the opening. Carefully controlled suction draws the cell into a tight fit and further trains of pressure pulses disrupt the membrane, leaving a patch of cell membrane spanning the pipette's lumen. A plastic jacket moulded around the pipette enables robotic handling. The machine holds up to six pipettes and each channel runs independently, so pipettes can be discarded as soon as the cell fails.

"Glass blowing enables us to be flexible in

the shape and geometry of the tips," says inventor Albrecht Lepple-Wienhues, founder and chief executive officer of flyion. This allows tailoring to suit different types of cells, and the new Flip-the-tip Large tips, which have a bowl at the base, enable the machine to monitor ligand-gated channels. "The bowl at the base gives us enough space to introduce a 130- μm diameter quartz needle," says Lepple-Wienhues. In flyion's standard tips, solution exchange takes about 60 seconds, but the new tips allow solutions to be puffed directly on to the cell through the quartz needle and give exchange rates of less than 50 milliseconds, while continuous recording is being carried out from each cell.

Patch-clamp economics

A report published in September by the Cambridge-based market-research consultancy HTStec makes interesting reading for those involved in the ion-channel industry. According to HTStec, the pharmaceutical and biotech market will spend around \$32 million in 2005 on automated patch-clamping machines. "We predict that sales will peak in 2006 at around 200 units a year," says HTStec director John Comley. In addition to this, the report estimates that for automated patch-clamping, labs spend around \$10 per data point for safety assessment and \$3.00 per data



flyion recording tip with a cell held in place at the end.

point in primary screening. Companies claim they would be more comfortable paying around \$6 and \$0.60 respectively. HTStec's survey of 66 companies and universities indicated that manual patch clamping was still the preferred option in assay development and safety assays such as hERG compliance, but automated patch clamping was the method of choice for secondary screening, lead optimization and early non-compliant hERG liability testing.

The highest possible throughput of some 3,000 data points a day is still far short of the 20,000 data points that respondents said they would like to get from a machine.

Many were looking forward to machines that measure ligand-gated channels much more cheaply. As these account for 29% of all ion channels studied, this is a potentially big market.

With genomics and proteomics creating a resurgence in cellular and systems research, there is every reason to believe that ion-channel research will become even more important in the coming decade.

Pete Moore is a science writer based near Bristol, UK.

Sigworth Laboratory

✉ info.med.yale.edu/cmphysiol/sigworth/HTStec

✉ www.htstec.com

SIGNALS OF DISEASE

Protein kinases are linked to numerous disease states, including cancer, arthritis, diabetes, cardiovascular diseases and neurological disorders. Gleevec from Novartis was the first compound active against a kinase (the Abl kinase) to be approved as a treatment — for certain gastrointestinal tumours and chronic myeloid leukaemia.

The market for kinases is large. "More than 25% of new drugs being developed today are based on kinase technology," says Jeff Linton, president of Upstate of Charlottesville, Virginia, which offers one of the largest collections of kinases. A flagship of Upstate's operation is its KinaseProfiler service, run from Dundee in Scotland. This provides quantitative characterization of compounds against an ever-expanding panel of human protein kinases in a direct radiometric assay. The panel currently



R&D Systems' Phospho-MAPK array tracks phosphorylated kinases.

contains around 230 kinases, almost 50% of the total number of human kinases in the genome. A new focus for Upstate is the addition of naturally occurring mutant kinases as they are identified.

Attention is also focusing on the newly emerging Gleevec-resistant mutants of Abl, and mutant forms

of other kinases including the epithelial growth factor (EGF) receptor, as these mutations can alter an inhibitor's efficacy. One of these mutations involves a single 'gatekeeper' amino acid. Mutations in this amino acid can prevent therapeutic compounds from binding effectively without affecting the enzyme's activity. "The search is on for successful inhibitors that are not sensitive to changes at the gatekeeper site," says Steve Davies, director of Upstate's drug discovery segment.

Upstate is helping this search by adding eight different mutant kinases to their portfolio, including ones for Kit, EGFR, Abl, Flt3 and p38/SAPK2a — and, says Davies, there are more in the pipeline.

A highly specific set of anti-kinase antibodies makes up R&D Systems' Proteome Profiler

Phospho-MAPK Array. This allows analysis of the phosphorylation status of 19 key signalling proteins, including members of all three major families of mitogen-activated protein kinases — the extracellular signal-regulated kinases, c-Jun N-terminal kinases, and the p38 kinases. These enzymes play essential roles in numerous signalling pathways that underlie cell function and disease.

Signalling pathway analysis products from Beckman Coulter of Fullerton, California, are also devoted to looking at intracellular activated (phosphorylated) kinases. One strength is that these reagents can be used on many different types of specimen including whole blood, and can resolve activated and inactivated kinases in whole blood cells, according to Michel Herbert, marketing manager for Beckman Coulter.

P.M.

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL
Ion channels and patch clamping			
Aurora Biomed	Ion Channel Reader (ICR) for voltage- and ligand-gated channels, assay services, cell lines	Vancouver, British Columbia	www.aurorabiomed.com
Aviva Biosciences	Equipment for electrophysiology, SEAL Microchips for high-throughput patch-clamp measurements	San Diego, California	www.avivabio.com
Cairn Research	Optopatch amplifier for patch-clamp recording, equipment for microscopy and electrophysiology	Faversham, UK	www.cairnweb.com
Collectricon	Dynaflow system for ion-channel drug discovery	Göteborg, Sweden	www.collectricon.com
ChanTest	Detection of intracellular ion-channel trafficking	Cleveland, Ohio	www.chantest.com
Cytomyx	Cell lines and cDNAs for kinases, receptors and ion channels	Cambridge, UK	www.cytomyx.com
flyion	Flyscreen high-throughput patch-clamp system	Tübingen, Germany	www.flyion.com
Invitrogen	Voltage sensor reagents, kits and reagents for genomics, proteomics, and cell biology	Carlsbad, California	www.invitrogen.com
Molecular Devices	PatchXpress and IonWorks for automated patch clamping, substrate finder for protein kinases	Sunnyvale, California	www.moleculardevices.com
Multi Channel Systems	Automated equipment for electrophysiology, <i>Xenopus</i> oocyte injection, ion-channel screening	Reutlingen, Germany	www.multichannelsystems.com
Nanion Technologies	Port-a-patch benchtop patch-clamping system for voltage- and ligand-gated ion channels	Munich, Germany	www.nanion.de
PerkinElmer Life Sciences	CellLux Fluorescence Cellular Screening Platform, instruments for cell biology	Boston, Massachusetts,	las.perkinelmer.com
Sophion Bioscience	QPatch 16 for automated patch clamping	Ballerup, Denmark	www.sophion.dk
Cell culture			
Cambrex	Poietics cell lines, gels, and products for molecular and cell biology research	Walkersville, Maryland	www.cambrex.com
Cytellect	Cell Express for development of cell lines with high levels of protein secretion	San Diego, California	www.cytellect.com
bSys	Development of cell lines for ion-channel screening	Basel, Switzerland	www.bsyst.ch
Cellular Dynamics International	Development of stem-cell derived cell lines for pharmaceutical screening	Madison, Wisconsin	
CruCell	Cell lines for the production of recombinant proteins and monoclonal antibodies	Leiden, The Netherlands	www.cruccell.com
Cureline	Tissue acquisition services	San Francisco, California	www.cureline.com
Cytogration	Automated high-throughput cell culture systems for in vitro screening	Rockville, Maryland	www.cytogration.com
EUGENEX Biotechnologies	Development of test cell lines	Taegervilen, Switzerland	www.eugenex.com
Gibco	Cell culture media and consumables	Carlsbad, California	www.invitrogen.com
Integra Biosciences	Equipment and consumables for sterilization, liquid handling, cell culture, sample storage	Baar, Switzerland	www.integra-biosciences.com
Kemp Biotechnologies	Contract services including cell culture and cell banking	Frederick, Maryland	www.kempbiotech.com
NanoDrop	Small sample spectrophotometry using fibre-optic technology	Wilmington, Delaware	www.nanodrop.com
Paragon Bioservices	Cell culture-based contract production and research services	Baltimore, Maryland	www.paragonbioservices.com
RTS Life Sciences	Automation for sample management, ultra-high-throughput screening and cell culture, and LIMS	Manchester, UK	www.rts-group.com
ScienCell Research Laboratories	Cell lines, including stem-cell lines, media, services	San Diego, California	www.sciencellonline.com
Stem Cell Sciences	Development of stem-cell lines	Edinburgh, UK	www.stemcellsciences.com
StemCell Technologies	Specialized media and cell-separation products for life-sciences research	Vacouver, British Columbia	www.stemcell.com
Vivascience	Ultrafiltration, protein purification, cell culture	Hannover, Germany	www.vivascience.com
Wave Biotech	Equipment for cell culture, Wave Bioreactor disposable cell culture systems	Bridgewater, New Jersey	www.wavebiotech.com
Xcellslz	Conditionally immortalized cell lines	Newcastle-on-Tyne, UK	www.xcellslz.com
Cell biology			
Active Motif	TransAM ELISA kits for transcription factor assays, antibodies, kits and reagents for cell fractionation	Carlsbad, California	www.activemotif.com
Advanced Targeting Systems	Targeted toxins, antibodies and custom services for the neurosciences	San Diego, California	www.atsbio.com
Applied BioPhysics	Automated cell-monitoring instruments measuring electrical impedance	Troy, New York	www.biophysics.com
Applied Cytometry Systems	Development and production of integrated cytometry systems	Sheffield, UK	www.appliedcytometry.us
BD Biosciences	FACS range of flow cytometers, antibody arrays, reagents and biochemicals for molecular biology	Franklin Lakes, New Jersey	www.bd.com
BioFX	Immunoreagents for biomedical research and diagnostics	Owing Mills, Maryland	www.biofx.com
BioGenex	Customized monoclonal and polyclonal antibodies, immunoassays, peptides	Berlin, Germany	www.biogenex.de
Biospherix	Instrumentation for hypoxia research	Redfield, New York	www.biospherix.com
Cayman Chemical	Immunoassays, antibodies, proteins, and small biological molecules for cell biology and neurobiology	Ann Arbor, Michigan	www.caymanchem.com
Eppendorf	Laboratory instrumentation and consumables for molecular and cell biology	Hamburg, Germany	www.eppendorf.com
Harvard Apparatus	Instruments and equipment for electrophysiology and cell biology	Holliston, Massachusetts	www.harvardapparatus.com
G Biosciences	Kits and reagents for proteome analysis, immunotechnology and molecular biology	St Louis, Missouri	www.genotech.com
GenScript	Kits and services for cell biology and molecular biology	Piscataway, New Jersey	www.genscript.com
Guava Technologies	Benchtop cytometers, live cell multi-parameter fluorescence assays	Hayward, California	www.guavatechnologies.com
Miltenyi Biotec	Magnetic cell-sorting technology	Cologne, Germany	www.miltenyibiotec.com
Trevigen	Antibodies, kits and reagents for cell biology	Gaithersburg, Maryland	www.trevigen.com
Warner Instruments	Electrophysiology and cell biology equipment and products	Hamden, Connecticut	www.warneronline.com
R&D Systems	Cytokines, interleukins, antibodies, flow cytometry kits, rat neural stem cells	Minneapolis, Minnesota	www.RnDSystems.com
Promega	Vectors, reagents and kits for genomics, proteomics and cellular analysis	Madison, Wisconsin	www.promega.com
Cell signalling			
Abcam	Antibodies, products and reagents for neuroscience, cell signaling, nuclear signaling, and chromatin studies	Cambridge, UK	www.abcam.com
Alexis Biochemicals	Reagents for cell biology, signal transduction, and molecular biology	Lausanne, Switzerland	www.alexis-corp.com
Assay Designs	MAP kinase ELISAs, Antibody Shop antibodies	Ann Arbor, Michigan	www.assaydesigns.com
Avanti Polar Lipids	Phospholipids and sphingolipids	Alabaster, Alabama	www.avantilipids.com
BD Biosciences	Catalogue and custom antibodies, kits for immunoassays	San Diego, California	www.bdbiosciences.com
Bender Medsystems	ELISA systems for research and in vitro diagnosis, instant ELISAs for cytokines	Vienna, Austria	www.bendermedsystems.com
BioImage	Cell-based assays and pathway screening for drug discovery	Soeborg, Denmark	www.bioimage.com
BIOMOL	Biochemicals for life-science research including signal transduction	Plymouth Meeting, Pennsylvania	www.biomol.com

COMPANY	PRODUCTS/ACTIVITY	LOCATION	URL
BioSource	Custom peptides, off-the-shelf and customized antibodies for signal transduction research	Camarillo, California	www.biosource.com
Calbiochem	Products for signal transduction, protein chemistry and glycobiology	San Diego, California	www.calbiochem.com
Cellomics	ArrayScan and KineticScan systems for automated screening for signal transduction pathways	Pittsburgh, Pennsylvania	www.cellomics.com
Cell Signaling Technology	Antibodies, proteins and kits for cell-signalling research	Beverly, Massachusetts	www.cellsignal.com
Chemicon	Catalogue and custom mono- and polyclonal antibodies, growth factors and cytokines	Temecula, California	www.chemicon.com
CIS bio International	HTRF (homogeneous time-resolved fluorescence)/TRACE reagents and kits for FRET assays	Bagnols-sur-Cèze, France	www.htrf-assays.com
cue Biotech	Minigene vectors for probing receptor-G-protein interactions	Evanston, Illinois	www.cuebiotech.com
DiscoverX	Cell-based assays for protein kinases and second messengers	Fremont, California	www.discoverx.com
GE Healthcare	Instruments, equipment and products for signal transduction assays and cellular assays	Little Chalfont, UK	www.amersham.co.uk
GloboZymes	Signal transduction enzymes, proteins and antibodies	Carlsbad California	www.globozymes.com
Hypomatrix	AntibodyArrays for signal transduction, cell cycle and apoptosis	Worcester, Massachusetts	www.hypomatrix.com
Intracellular Imaging	Imaging systems for calcium signaling	Cincinnati, Ohio	www.intracellular.com
JPT Peptide Technologies	Kinase substrate identification service, peptide substrates	Berlin, Germany	www.jerini.com
Mabtech	Antibodies for cytokine detection	Nacka Strand, Sweden	www.mabtech.comcom
New England Biolabs	Enzyme technology, cell-signaling reagents and antibodies	Beverly, Massachusetts	www.neb.com
OriGene Technologies	Human full-length cDNA clones, including GPCRs, protein kinases, and nuclear hormone receptors	Rockville, Maryland	www.origene.com
Panomics	TranSignal Human Cytokine Antibody Arrays	Redwood City, California	www.panomics.com
Pepscan Systems	Custom and catalog PepChip kinase substrate peptide arrays for kinase profiling, services	Lelystad, The Netherlands	www.pepscan.com
ProKinase	Validation of new protein kinases as drug targets, recombinant protein kinases	Freiburg, Germany	www.prokinase.com
Ray Biotech	Cytokine antibody arrays	Norcross, Georgia	www.raybiotech.com
Rockland	Antibodies for cell-signalling research, immunochemicals	Gilbertsville, Pennsylvania	www.rockland-inc.com
Sigma-Aldrich	Reagents for life-sciences research, Panorama antibody arrays for cell-signalling proteins	St Louis, Missouri	www.sigmaaldrich.com
Stressgen	Reagents and kits for cellular stress, apoptosis, oxidative stress and neurobiology research	Victoria, British Columbia	www.stressgen.com
SuperArray Bioscience	GEArray Signal Transduction Gene Arrays, application-specific cDNA	Frederick, Maryland	www.superarray.com
Tocris Cookson	Chemicals for life-science research, including signal transduction agents	Avonmouth, UK	www.tocris.com
Upstate Group	KinaseProfiler specificity testing, IC50Profiler Express, reagents for cell-signalling research	Charlottesville, Virginia	www.upstate.com

Imaging for cell biology

Applied Scientific Instrumentation	Automated systems for microscopy and fluorescence screening	Eugene, Oregon	www.asiimaging.com
Carl Zeiss	Microscopes, laser-scanning microscopes, fluorescence correlation spectroscopy, confocal imaging	Gottingen, Germany	www.zeiss.com
Hamamatsu Photonic Systems	Digital imaging systems and cameras	Hamamatsu City, Japan	www.hamamatsu.com
Intelligent Imaging Innovations	Digital microscopy workstations and imaging software	Denver, Colorado	www.intelligent-imaging.com
Kodak	Digital imaging equipment	Rochester, New York	www.kodak.com
LaVision Biotech	TriMScope multifocal multiphoton laser-scanning microscope	Bielefeld, Germany	www.lavisionbiotech.de
Leica Microsystems	DM digital microscope range, microscopes and microscope systems	Wetzlar, Germany	www.leica-microsystems.com
Nikon Instruments	Microscopes, COOLSCOPE digital microscope with screen	Kanagawa, Japan	www.nikon-instruments.jp/eng/
Olympus	Microscopes, Fluoview confocal laser-scanning microscopes, cell range of imaging stations	Hamburg, Germany	www.olympus-europa.com
Optronics	Digital microscope cameras, image-analysis software	Goleta, California	www.optronics.com
Polaroid	Digital imaging systems for microscopes	Waltham, Massachusetts	www.polaroid.com
Princeton Instruments	Equipment for imaging and spectroscopy	Trenton, New Jersey	www.princetoninstruments.com
QImaging	High-performance digital cameras and RGB filters for microscopy	Burnaby, British Columbia	www.qimaging.com
Richardson Technologies	RTM 3.0 microscope system for live-cell imaging, high-resolution field microscope	Bolton, Ontario	www.richardson-tech.com
Varian	Spectrophotometers, IR microscopes and imaging systems	Palo Alto, California	www.varianinc.com
VayTek	Digital imaging systems, cameras, imaging software	Fairfield, Iowa	www.waytek.com
Vision Engineering	Microscope manufacturers, microscope accessories	Woking, UK	www.visioneng.com

General

Abgent	Catalogue and customized antibodies, proteins and peptides, protein expression services	San Diego, California	www.abgent.com
Agilent	Instrumentation, kits and consumables for the life sciences	Palo Alto, California	www.agilent.com
Applied Biosystems	DNA sequencers, workstations and automated instruments for genomics and proteomics	Foster City, California	home.appliedbiosystems.com
Beckman Coulter	Automated tools for molecular biology, biochemistry, genomics and proteomics	Fullerton, California	www.beckmancoulter.com
Bio-Rad	Products, instruments and software for life-sciences research	Hercules, California	www.bio-rad.com
BMG LABTECH	Microplate and array readers and handling systems	Offenburg, Germany	www.bmg-labtechnologies.com
Cambridge Research Biochemicals	Peptide synthesis and modification, antibody production, ELISA	Billingham, UK	www.crb.gb.com
Charles Rivers Laboratories	Services for biopharmaceutical and preclinical research	Wilmington, Massachusetts	www.criver.com
Daivids Biotechnologie	Custom antibody development	Regensburg, Germany	www.dabio.de
Gilson	Liquid-handling and pipetting instruments	Middleton, Wisconsin	www.gilson.com
Millipore	Automated equipment for molecular biology, biochemistry, genomics and proteomics	Bedford, Massachusetts	www.millipore.com
Nalge Nunc International	Labware	Rochester, New York	www.nalgenunc.com
Pierce Biotechnology	Components and consumables for cell and molecular biology	Rockford, Illinois	www.piercenet.com
Pro Scientific	Laboratory equipment, centrifuges	Oxford, Connecticut	www.proscientific.com
Qiagen	BioRobot multifunctional laboratory workstation, kits and instruments	Valencia, California	www.qiagen.com
Roche Applied Science	Reagents and kits for molecular biology, functional genomics and proteomics research	Indianapolis, Indiana	www.roche-applied-science.com
Stratagene	Tools and reagents for molecular biology, genomics and proteomics	La Jolla, California	www.stratagene.com
TaKaRa Bio	Reagents, kits and custom services for life-sciences research	Shiga, Japan	www.takara.bio.com
Wako Chemicals	Specialty chemicals, bioproducts and clinical diagnostic reagents	Richmond, Virginia	www.wakousa.com

● Advertiser

-  FOCUS
-  SPOTLIGHT
-  RECRUITMENT
-  ANNOUNCEMENTS
-  EVENTS

naturejobs

Physics in crisis?

Physics is failing to attract fresh blood in England and Wales. A survey published last month by the University of Buckingham reveals that the number of pupils taking A-level physics has fallen by 38% since 1990. And in the same period, the proportion of new physics teachers dropped from one-third of all science teachers to just 13%.

The results ring alarm bells as far as the supply pipeline for physicists is concerned. Where will the next generation of physics researchers in England and Wales come from?

The United States has faced a similar issue — and has managed to tackle it with some degree of success. Between 1987 and 1997, the proportion of high-school students taking physics rose from 20% to 28%. This was thanks to several initiatives, including a broadening of the school curricula so that the regular physics course was supplemented by courses on application, theory and basic concepts, as well as university-level courses for more advanced students. In addition, the schools did more outreach to female students and engaged with the scientific community.

One programme has high-school students working with

research scientists at the Laser Teaching Center in the State University of New York, Stony Brook. It offers a chance to do some hands-on work together with laser and optical physicists, and puts high-school students shoulder-to-shoulder with PhDs.

But despite this sort of outreach, and the increase in high-school science enrolment, the number of US students taking graduate-level physics has declined steadily in recent years. In 2003, fewer than 500 US citizens earned physics PhDs, the lowest number since the early 1960s. The result has been an increased dependence on foreign-born physicists.

Perhaps what is needed in the United States, England and Wales is to create more jobs in the field and to communicate the excitement that comes along with doing physics, whatever the setting.



Paul Smaglik, *Naturejobs* editor

CONTACTS

Publisher: Ben Crowe
Editor: Paul Smaglik
Assistant Editor: Corie Lok

European Head Office, London
 The Macmillan Building, 4 Crinan Street
 London N1 9XW, UK
 Tel: +44 (0) 20 7843 4961
 Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com

Naturejobs Sales Director:
 Nevin Bayoumi (4978)
European Sales Manager:
 Andy Douglas (4975)

Natureevents: Sille Opstrup (4994)
UK/RoW/Ireland/Italy:
 Nils Moeller (4953)
 Irene Viglia-Atton (4944)
Scandinavia/Spain/Portugal:
 Evelina Rubio Håkansson (4973)
France/Switzerland/Belgium:
 Amelie Pequignot (4974)
Germany/Austria/The Netherlands:
 Reya Silao (4970)

Advertising Production Manager:
 Billie Franklin
 To send materials use London
 address above.
 Tel: +44 (0) 20 7843 4814

Fax: +44 (0) 20 7843 4996
 e-mail: naturejobs@nature.com
Naturejobs web development: Tom Hancock
Naturejobs online production: Niamh Shields

European Satellite Office
Germany/Austria/
The Netherlands:
 Patrick Phelan
 Tel: +49 89 54 90 57 11
 Fax: +49 89 54 90 57 20
 e-mail: p.phelan@nature.com

US Head Office, New York
 345 Park Avenue South,
 10th Floor, New York,

NY 10010-1707
 Tel: +1 800 989 7718
 Fax: +1 800 989 7103
 e-mail: naturejobs@natureny.com

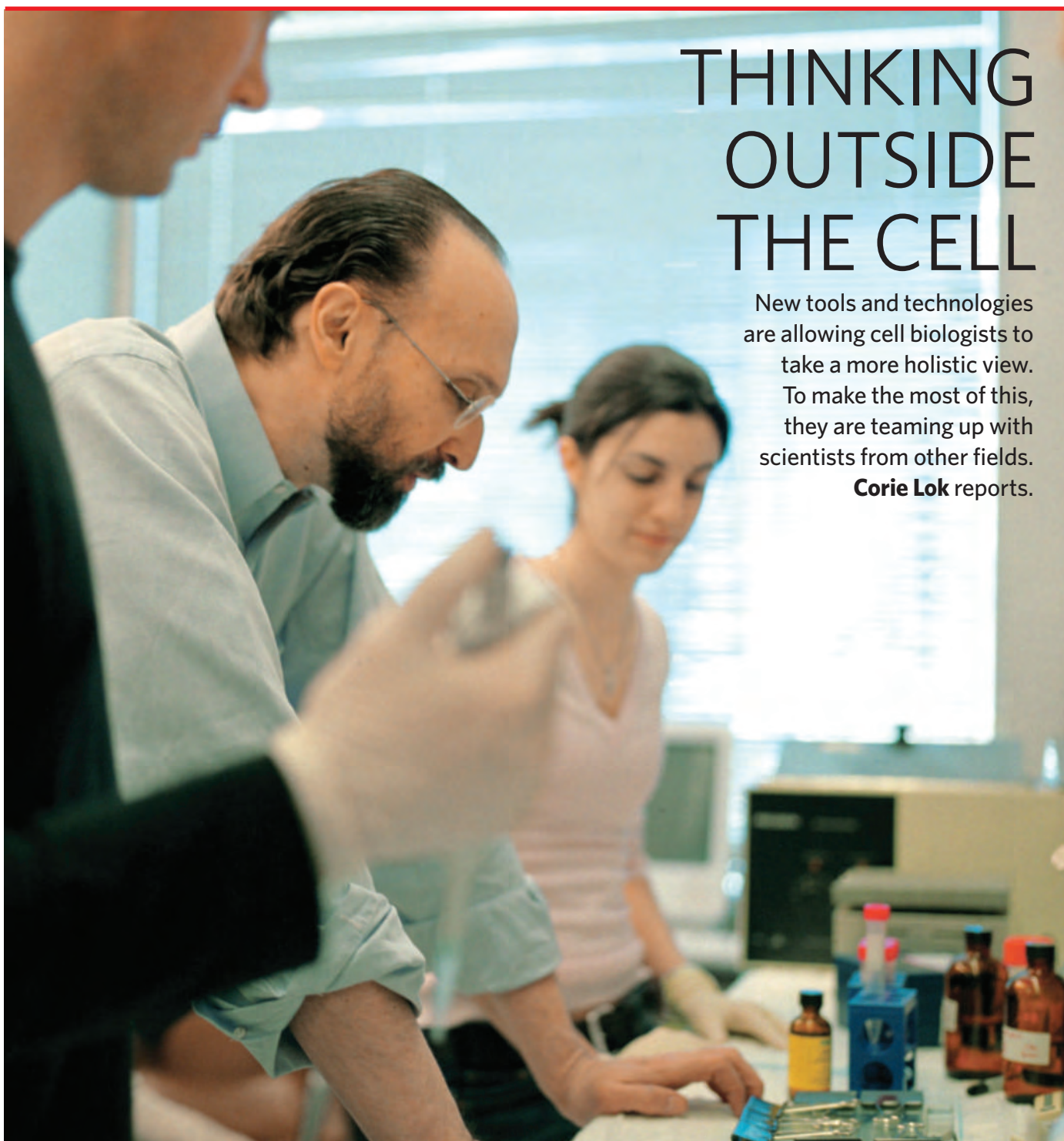
US Sales Manager: Peter Bless

Japan Head Office, Tokyo
 Chiyoda Building,
 2-37 Ichigayatamachi,
 Shinjuku-ku,
 Tokyo 162-0843
 Tel: +81 3 3267 8751
 Fax: +81 3 3267 8746
Asia-Pacific Sales Director: Rinoko Asami
 e-mail: rasami@naturejpn.com

THINKING OUTSIDE THE CELL

New tools and technologies are allowing cell biologists to take a more holistic view. To make the most of this, they are teaming up with scientists from other fields.

Corie Lok reports.



For the past 20 years, Joan Massagué has been studying the growth factors and signalling pathways involved in cell proliferation and regulation. He has spent most of those years at the Memorial Sloan-Kettering Cancer Center in New York, but his research was still in basic cell biology.

"It was called cancer research because it was being done in the name of cancer, but it wasn't really research on cancer," says Massagué. Now that has changed.

About three or four years ago, Massagué began a project focused on the mechanisms underlying metastasis. Little is known about how tumour cells spread to other tissues and organs. But by then, Massagué felt he and his lab had accumulated enough knowledge and tools through years of basic research to

"We want students to acquire a strong clinical perspective and to understand patient-oriented research."

— Kenneth Mariani

start on the problem. "If I'd made this decision five or ten years ago, I would have gone nowhere," he says. "It's still very difficult and it costs a lot of money, but it can be done now." Earlier this year, his lab published a paper in *Nature* identifying genes that mediate the spread of breast cancer tumours to the lungs (A. J. Minn *et al. Nature* **436**, 518–524; 2005).

Because of the new metastasis programme, Massagué has been recruiting more MD/PhD students and MDs as postdocs. Even those trained as basic biologists want to work on the more applied projects. He receives three times more applications for the metastasis projects than for his lab's ongoing basic-science projects.

Massagué's foray into such 'translational' research shows how cell biology is stretching beyond its

P. FETTERS

traditional borders. With the wealth of knowledge in basic cell biology and the rapid advance in powerful tools and technologies, cell biologists can now ask increasingly complex questions and find applications for the answers. After so many years of deconstructing the cell and its parts, researchers are putting the pieces together. To do so, they are turning to people outside the field — not just clinicians for translation help, but also physicists and engineers to build mathematical models illustrating the dynamics of cellular activities.

Even with this move towards interdisciplinary work, cell biologists emphasize that young researchers more interested in fundamental questions in biology are still in demand. Indeed, about half of Massagué's lab still conducts basic research. But now there is a need for a new kind of cell biologist: one who can, for example, work with human tissue samples and understand a disease process in patients; one who can think quantitatively and communicate with a mathematician; one who can co-author a paper with a physicist.

"What I'm looking for in my lab is interdisciplinary thinking; people who are fearless in pursuing new systems or new technologies," says Randall Moon, a biologist at the University of Washington in Seattle.

A trip to the hospital

Throughout the Sloan-Kettering Cancer Center, biology is shifting towards the clinic. The new Gerstner Sloan-Kettering Graduate School of Biomedical Sciences is recruiting its first class of graduate students, about a dozen, to begin in autumn 2006. The programme will train students to be strong bench scientists. "We also want them to acquire a strong clinical perspective and to understand patient-oriented research," says Kenneth Mariani, chair of the molecular-biology programme at the cancer centre and dean of the new school.

The students will conduct full-time lab research, taking a cancer-biology class and working in different labs during their first year before selecting a thesis adviser and launching into their research project. But they can also choose a clinical mentor from the affiliated Memorial Hospital, which could enable them to attend clinical seminars and conferences.

At Stanford University's medical school, another PhD programme designed to bring basic biology and medicine closer will start in September 2006. The first class — six PhD students from Stanford's various biology departments — will take medical-school classes during the first year and a half. They will also take graduate classes and do lab rotations like the other PhD students. In the second year, they will begin their thesis work as usual. By the end, they will graduate with both a PhD and an MSc in medicine. Ben Barres, a neurobiologist who is spearheading the programme, says it is geared towards students who are interested in translational research, but who don't want to spend seven or more years in a MD/PhD programme.

A handful of other US medical schools, including Harvard, have established similar programmes to bridge the divide between basic and clinical sciences.

In Scotland, the Research Institute for Medical Cell Biology at the University of Edinburgh opened this summer and houses 600 bench researchers. It is located next to a new 900-bed hospital and teaching facility to encourage a 'bench to bedside' approach to biological research.

"Clinical research has had an image of being simple,"



Building bridges between disciplines: Kenneth Mariani (opposite), biologist Raymond Deshaies (top), cancer researcher Joan Massagué (middle) and mathematician Alex Mogilner.

says Alan Hall, director of the UK Medical Research Council's Laboratory for Molecular Cell Biology at University College London. But that is changing, he says. "I think there's a huge need for cell biologists in translational research," agrees Barres.

Cell biology is branching out not just into the clinical realm, but also into mathematics. Now that cell biologists have sophisticated tools to generate a plethora of data, they need mathematical models to put this information together in a meaningful and comprehensive way. And that means collaborating with engineers, physicists and mathematicians. Modelling is becoming more mainstream in cell biology, thanks to the growing sophistication of the models and the increasing amount of high-quality data to feed into them.

"The rate at which modelling is becoming acceptable has been amazing in the past few years," says Alex Mogilner, a professor in the maths department and the Center for Genetics and Development at the University of California, Davis.

Accommodating complexity

To foster this interdisciplinary work, Cornell University in Ithaca, New York, is building a cell and molecular biology institute, due to open in 2007. The building will house a dozen newly hired cell biologists, nine junior and three senior investigators, plus several other researchers from physics, nanotechnology and other physical sciences in which Cornell has traditionally been strong. "We've now come to appreciate the complexity of the questions we're asking," says Cornell biochemist Rick Cerione, who heads the committee seeking a director for the institute. "We need to bring multiple disciplines to bear."

At the Center for Integrative Molecular Biosciences at the Scripps Research Institute in La Jolla, California, engineer Gaudenz Danuser has teamed up with cell biologist Clare Waterman-Storer. Each mathematician in his lab is matched up with a biologist in Waterman-Storer's lab to work on problems in pairs. At Davis, Mogilner runs a lab jointly with biologist Jonathan Scholey, where they pool resources and students.

It is not clear what the best path is to computational cell biology, or even if there is one. Students can learn both biology and maths in the biological-engineering and systems-biology departments that are cropping up at universities such as the Massachusetts Institute of Technology and Harvard. Danuser says it is better for students to specialize in one or the other while being open to working with people outside their field. One option for a young cell biologist is to do a postdoctoral fellowship in a lab that does a lot of modelling or that collaborates with a lab such as Danuser's.

All those engaging in interdisciplinary research need to be open-minded, flexible and motivated, says Ivan Dikic, a biochemist at the Goethe University School of Medicine in Frankfurt, Germany.

Strengthening the quantitative background would be a potentially useful thing for young cell biologists, says Raymond Deshaies, a biologist at the California Institute of Technology in Pasadena. He lists physics, maths, statistics and even computer programming as important areas. "Even if you don't use them yourself on a day-to-day basis, you might end up collaborating with people from these fields," says Deshaies. "At least you'll be able to talk to them."

Corie Lok is assistant editor of Naturejobs.

The Albian message

A blast from the past.

Oliver Morton

To: Eva P.

From: Stefan K.

Re: Sample handling facility

March 4, 2047

I thought I ought to put into writing my concerns over the sample-return facility for Odyssey. I think that relying on the mothballed Mars Sample Return lab at Ames is dangerously complacent. It is simply not flexible enough, or big enough, for what I think we should be expecting.

I appreciate that I am in a minority on this, and that the consensus is that we will be dealing with non-biological artefacts. And I don't want to sound like the people from AstraRoche slipped some egopietin into my drink during that trip to Stockholm last November. But my minority views have been pretty well borne out throughout this whole story. Back when Suzy and Sean had more or less convinced the world that the trinity sequences in the Albian message referred to some sort of mathematico-philosophical doctrine — possibly based on an analogy to the aliens' purported trisexual reproductive system — and everyone in SETI was taking a crash course in genome analysis, I had to pull in every favour I was owed to get the Square Kilometre Array used as a planetary radar and scanned over the Trojan asteroids. If I hadn't done that we wouldn't even know about the Pyramid, let alone be sending Odyssey there.

I'm not claiming I understand the Albians' minds better than anyone else; I haven't got any more of the message in my DNA than anyone else has. And it's always been my position that we should read as little into that message as possible. I remain convinced that looking for descriptions of their philosophy or lifestyle or even provenance is pointless. The more I look at the increasingly meaningless analyses that the increasingly intelligent AIs produce, the more I think that the variations between phyla are effectively random and that the message from the aliens tells us almost nothing except that there's a radar-reflecting tetrahedron $\pi/3$ behind Jupiter that they think we may find interesting.

Everyone assumes that if it hadn't been for the parts of the message lost in the K/T the 'residual variant sequences' would be seen to add up to some great big life-the-Universe-and-everything revelation. And

because they think such a revelation once existed, they expect to see it carved into the palladium walls of the Pyramid. But if the aliens who visited Earth, and left their messages in the genomes of more or less everything on the planet, had wanted to tell us something more about themselves, they could have made the messages a lot bigger and built in more redundancy across phylum space; there's no shortage of junk DNA to write on. The point is, they didn't choose to leave big messages — just a simple signpost.

The reason I was able to get the SKA people to find the Pyramid was that they knew I'd thought about SETI a lot. But these days people tend to forget that I was always something of a sceptic. What could a bunch of aliens tell us about themselves, or the Universe, that would matter? Especially if, like the Albians, they sent, or rather left, the message 100 million years ago? Well, in the case of the Albians, there's one type of knowledge they could be fairly sure that anyone who eventually evolved sequencing technology on Earth pretty much had to be interested in. And it's something that, by definition, is too big to fit into the spare bits of a genome.

I appreciate that everyone on the project now has a lot of faith in what we can do on the fly, especially in terms of recording and analysing information. I'll admit that when we started I really didn't think that the lost

craft of human spaceflight would be so easy to reinvent. It still strikes me as remarkable that none of us realized how much could be achieved by leaving a technical problem to one side and concentrating on other things for a few decades before coming back to it with new technologies. But the problem with the sample-return facility won't just be one of technology. It's going to be one of size.

You see, extinctions aren't the noise in the message. They're the reason for the message. The one thing the Albians knew they could do for whoever would end up reading their message was store up some of the biodiversity that would inevitably be whittled away over time. When Odyssey gets to the Trojan Pyramid, I don't expect it to find any more information about the Albians than we have already. I do expect a biosphere's worth of well-preserved biological samples from the mid-Cretaceous. Not just genomes, but whole samples. Sudarat and her boys are going to come home with a hold full of early angiosperms and dinosaur eggs. We need to be ready. ■ Oliver Morton is a contributing editor at *Wired* and the author of *Mapping Mars: Science, Imagination and the Birth of a World* (2002) and the forthcoming *Eating the Sun: How Plants Power the Planet*. He does not normally knowingly commit fiction. He took over as *Nature's* Chief News and Features Editor last month.



JACEY